普通高等教育"十一五"国家级规划教材
卫生部"十一五"规划教材
全国高等医药教材建设研究会规划教材
全国高等学校教材
供8年制及7年制临床医学等专业用

第2版

医学遗传学

Medical Genetics

主　编　陈　竺
副主编　陆振虞　傅松滨

人民卫生出版社
PEOPLE'S MEDICAL PUBLISHING HOUSE

全国高等医药教材建设研究会规划教材

全国高等学校教材

供**8年制**及**7年制**临床医学等专业用

医学遗传学

Medical Genetics

- * 1.《细胞生物学》第2版（含光盘）
- * 2.《系统解剖学》第2版（含光盘）
- * 3.《局部解剖学》第2版（含光盘）
- * 4.《组织学与胚胎学》第2版（含光盘）
- * 5.《生物化学与分子生物学》第2版（含光盘）
- * 6.《生理学》第2版（含光盘）
- * 7.《医学微生物学》第2版（含光盘）
- * 8.《人体寄生虫学》第2版（含光盘）
- * 9.**《医学遗传学》第2版（含光盘）**
- *10.《医学免疫学》第2版
- *11.《病理学》第2版（含光盘）
- *12《病理生理学》第2版（含光盘）
- *13.《药理学》第2版（含光盘）
- *14.《临床诊断学》第2版（含光盘）
- *15.《实验诊断学》第2版（含光盘）
- *16.《医学影像学》第2版（含光盘）
- *17.《内科学》第2版（含光盘）
- *18.《外科学》第2版（含光盘）
- *19.《妇产科学》第2版（含光盘）
- *20.《儿科学》第2版（含光盘）
- *21.《感染病学》第2版（含光盘）
- *22.《神经病学》第2版（含光盘）
- *23.《精神病学》第2版（含光盘）
- *24.《眼科学》第2版（含光盘）
- *25.《耳鼻咽喉头颈外科学》第2版
- *26.《核医学》第2版（含光盘）
- *27.《预防医学》第2版（含光盘）
- *28.《医学心理学》第2版（含光盘）
- 29.《医学统计学》第2版（含光盘）
- *30.《循证医学》第2版（含光盘）
- *31.《医学文献信息检索》第2版（含光盘）
- 32.《临床流行病学》（含光盘）
- 33.《肿瘤学》
- 34.《生物信息学》（含光盘）
- 35.《实验动物学》（含光盘）
- 36.《医学科学研究导论》
- 37.《医学伦理学》（含光盘）

注：全套书均为卫生部“十一五”规划教材，
画 * 者为普通高等教育“十一五”国家级规划教材

策划编辑 杨 晋
责任编辑 杨 晋 马 瑛
封面设计 郭 淼
版式设计 李秋𠅘

人民卫生出版社网站：
门户网：www.pmph.com 出版物查询、网上书店
卫人网：www.ipmph.com 护士、医师、药师、中医师、卫生资格考试培训

ISBN 978-7-117-12817-9

定价(含光盘)：56.00元

- 普通高等教育“十一五”国家级规划教材
- 卫生部“十一五”规划教材
- 全国高等医药教材建设研究会规划教材
- 全国高等学校教材
- 供8年制及7年制临床医学等专业用

第2版

医学遗传学

Medical Genetics

主　编　陈　竺

副主编　陆振虞　傅松滨

编　者（以章节出现先后为序）

陈　竺（上海交通大学医学院）

吴白燕（北京大学医学部）

罗泽伟（复旦大学）

傅松滨（哈尔滨医科大学）

彭鲁英（同济大学生命科学学部）

赵彦艳（中国医科大学）

左　伋（复旦大学上海医学院）

陆振虞（上海交通大学医学院）

孙树汉（第二军医大学）

张思仲（四川大学华西医学中心）

秘　书　倪萦音（上海交通大学医学院）

人民卫生出版社

图书在版编目（CIP）数据

医学遗传学/陈竺主编．—2版．—北京：人民卫生出版社，2010.7

ISBN 978-7-117-12817-9

Ⅰ.①医… Ⅱ.①陈… Ⅲ.①医学遗传学-高等学校-教材 Ⅳ.①R394

中国版本图书馆CIP数据核字(2010)第072818号

门户网：www.pmph.com	出版物查询、网上书店
卫人网：www.ipmph.com	护士、医师、药师、中医师、卫生资格考试培训

医 学 遗 传 学

第 2 版

主　　编：陈　竺

出版发行：人民卫生出版社（中继线 010-59780011）

地　　址：北京市朝阳区潘家园南里19号

邮　　编：100021

E-mail：pmph @ pmph.com

购书热线：010-67605754　010-65264830

010-59787586　010-59787592

印　　刷：中国农业出版社印刷厂

经　　销：新华书店

开　　本：850×1168　1/16　**印张：**19

字　　数：562千字

版　　次：2005年8月第1版　2010年7月第2版第6次印刷

标准书号：ISBN 978-7-117-12817-9/R·12818

定价（含光盘）：56.00元

打击盗版举报电话：010-59787491　E-mail：WQ @ pmph.com

（凡属印装质量问题请与本社销售中心联系退换）

第二版出版说明

全国高等学校八年制临床医学专业规划教材自2005年出版以来，得到了教育部、卫生部等主管部门的认可，以及医学院校广大师生的好评。为了进一步满足教学改革与实践不断推进，以及医学科学不断发展的需要，全国高等医药教材建设研究会和卫生部教材办公室在吴阶平、裘法祖、吴孟超、陈灏珠和刘德培院士等的亲切关怀和支持下于2009年启动了该套教材第二轮的修订工作。

第二轮修订过程中仍坚持"精品战略，质量第一的原则，从精英教育的特点、医学模式的转变、信息社会的发展、国内外教材的对比等角度出发，在注重'三基'、'五性'的基础上，从内容到形式都'更新'、'更深'、'更精'，为培养高素质、高水平、富有临床实践和科学创新能力的医学博士服务"的编写宗旨，并根据使用过程中的反馈意见与建议，在第一轮的基础上力求做到：学科体系更加完善，增加了《临床流行病学》、《肿瘤学》、《生物信息学》、《实验动物学》、《医学科学研究导论》和《医学伦理学》；相关学科的交叉与协调更为完善，比如《生物化学》与《医学分子生物学》合并为《生物化学与分子生物学》；内容的选材与框架体系的设计更加注重启发性，强调学生创新能力的培养，并适当给学生留下了思维分析、判断、探索的空间；教材的配套更加健全；装帧设计更为精美。

该套书在修订过程中，得到了广大医学院校的大力支持，作者均来自各学科临床、科研、教学第一线，具有丰富临床、教学、科研和写作经验的优秀专家，作者队伍覆盖了目前国内所有开办临床医学专业八年制及七年制的院校。

修订后的第二版仍以全国高等学校临床医学专业八年制及七年制师生为主要目标读者，并可作为研究生、住院医师等相关人员的参考用书。

全套教材共37种，其中36种于2010年8月出版，1种将于2010年年底出版。

八年制教材目录

*1.《细胞生物学》 第2版（含光盘）	主　编 副主编	杨　恬 左　伋　刘艳平
*2.《系统解剖学》 第2版（含光盘）	主　编 副主编	柏树令 应大君　丁文龙　崔益群
*3.《局部解剖学》 第2版（含光盘）	主　编 副主编	王怀经　张绍祥 张雅芳　胡海涛
*4.《组织学与胚胎学》 第2版（含光盘）	主　编 副主编	高英茂　李　和 李继承　陈晓蓉
*5.《生物化学与分子生物学》 第2版（含光盘）	主　编 副主编	贾弘禔　冯作化 屈　伸　药立波　方定志　冯　涛
*6.《生理学》 第2版（含光盘）	主　编 副主编	姚　泰 曹济民　樊小力　王庭槐
*7.《医学微生物学》 第2版（含光盘）	主　编 副主编	贾文祥 陈锦英　江丽芳　黄　敏
*8.《人体寄生虫学》 第2版（含光盘）	主　编 副主编	詹希美 诸欣平　刘佩梅
*9.《医学遗传学》 第2版（含光盘）	主　编 副主编	陈　竺 陆振虞　傅松滨
*10.《医学免疫学》 第2版	主　编 副主编	何　维 曹雪涛　熊思东
*11.《病理学》 第2版（含光盘）	主　编 副主编	陈　杰　李甘地 文继舫　来茂德　孙保存
*12.《病理生理学》 第2版（含光盘）	主　编 副主编	李桂源 吴伟康　欧阳静萍
*13.《药理学》 第2版（含光盘）	主　编 副主编	杨世杰 杨宝峰　颜光美　臧伟进
*14.《临床诊断学》 第2版（含光盘）	主　编 副主编	欧阳钦 吴汉妮　刘成玉
*15.《实验诊断学》 第2版（含光盘）	主　编 副主编	王鸿利 尚　红　王兰兰
*16.《医学影像学》 第2版（含光盘）	主　编 副主编	金征宇 冯敢生　冯晓源
*17.《内科学》 第2版（含光盘）	主　编 副主编	王吉耀 廖二元　黄从新　华　琦
*18.《外科学》 第2版（含光盘）	主　编 副主编	陈孝平 石应康　邱贵兴　杨连粤

*19.《妇产科学》 第2版（含光盘）	主　编 副主编	丰有吉　沈　铿 马　丁　孔北华　李　力
*20.《儿科学》 第2版（含光盘）	主　编 副主编	薛辛东 杜立中　毛　萌
*21.《感染病学》 第2版（含光盘）	主　编 副主编	王宇明 施光峰　宁　琴　李　刚
*22.《神经病学》 第2版（含光盘）	主　编 副主编	吴　江 贾建平　崔丽英
*23.《精神病学》 第2版（含光盘）	主　编 副主编	江开达 于　欣　李凌江　王高华
*24.《眼科学》 第2版（含光盘）	主　编 副主编	葛　坚 赵家良　黎晓新
*25.《耳鼻咽喉头颈外科学》 第2版	主　编 副主编	孔维佳 周　梁　许　庚　王斌全　唐安洲
*26.《核医学》 第2版（含光盘）	主　编 副主编	张永学　黄　钢 匡安仁　李亚明
*27.《预防医学》 第2版（含光盘）	主　编 副主编	孙贵范 凌文华　孙志伟　姚　华
*28.《医学心理学》 第2版（含光盘）	主　编 副主编	姜乾金 马　辛　林大熙　张　宁
29.《医学统计学》 第2版（含光盘）	主　编 副主编	颜　虹 徐勇勇　赵耐青
*30.《循证医学》 第2版（含光盘）	主　编 副主编	王家良 詹思延　许能锋　康德英
*31.《医学文献信息检索》 第2版（含光盘）	主　编 副主编	罗爱静 马　路　于双成
32.《临床流行病学》 （含光盘）	主　编 副主编	李立明 詹思延　谭红专
33.《肿瘤学》	主　编 副主编	郝希山　魏于全 赫　捷　周云峰
34.《生物信息学》 （含光盘）	主　编 副主编	李　霞 李亦学　廖　飞
35.《实验动物学》 （含光盘）	主　编 副主编	秦　川 张连峰　魏　泓　顾为望　王　钜
36.《医学科学研究导论》	主　编 副主编	詹启敏　赵仲堂 刘　佳　刘　强
37.《医学伦理学》 （含光盘）	主　编 副主编	王明旭 尹　梅　严金海

注：全套书均为卫生部“十一五”规划教材，画*者为普通高等教育“十一五”国家级规划教材

八年制教材再版序言

五年来，在大家的热情呵护下，我们共同见证了八年制临床医学教材——这个新生命的诞生与茁壮成长。如今，第二版教材与大家见面，怀纳第一版之精华而不张扬，吞吐众学者之智慧而不狂放，正如医学精英人才所应具备的气质与神韵。在继承中发展，新生才能越发耀眼；切时代之脉搏，思维才能永领潮头。第二版教材已然跨入新的成长阶段，心中唯觉欣喜和慰藉。

回想第一版教材面世之后，得到了各方众多好评，这充分说明了：这套教材将生命科学信息化、网络化以及学科高度交叉、渗透的特点融于一身，同时切合了环境-社会-心理-工程-生物医学模式的转变，诠释了以人为本、协调发展的战略思想。另外，编委构成的权威性和代表性、内容选择、编排体系、印刷装帧质量等，令广大师生耳目一新，爱不释卷。诚然，第一版教材也并非十全十美，比如有的学科仍以介绍知识为主，启发性不强，对学生难以起到点石成金、抛砖引玉的作用，不利于学生创新思维能力的培养；有的学科、章节之间有重复现象，略显冗余，不够干练。另外，随着学科的进展，部分疾病的临床分类、治疗等内容已略显滞后，亟待最新的研究成果加入其中，充实完善。

鉴此，第一版教材的修订工作便提上日程。此次修订，比当初第一版的编纂过程更为艰辛和严谨，从编者的谨慎遴选到教材内容的反复推敲、字斟句酌，可谓精益求精、力臻完美，经过数轮探讨、分析、总结、归纳、整理，第二版教材终于更富于内涵、更具有生命力地与广大师生们见面了。

“精英出精品，精品育精英”是第二版教材在修订之初就一直恪守的理念。主编、副主编与编委们均是各领域内的医学知名专家学者，不仅著作立身，更是德高为范。在教材的编写过程中，他们将从医执教中积累的宝贵经验、体会以及医学精英的特质潜移默化地融入到教材当中。同时，在主编负责制的前提下，主编、副主编负责全书的系统规划，编委会构成团结战斗的团队，各位专家群策群力、扬长补短、集思广益、查漏补缺，为教材的高标准、高质量的修订出版打下了坚实的基础。

注重医学学科内涵的延伸与发展，同时兼顾学科的交叉与融合是第二版教材的一大亮点。此次修订不仅在第一版的基础上增加了《临床流行病学》、《肿瘤学》、《生物信息学》、《实验动物学》、《医学科学研究导论》和《医学伦理学》，同时还合并了《生物化学》与《医学分子生物学》。通过主编顶层设计，相邻学科主编、副主编协调与磋商，互审编写提纲，以及交叉互审稿件等措施，相当程度上实现了突出中心、合理交叉、避免简单重复的要求。

强调启发性以及创新意识、创新思维和创新能力的培养是第二版教材的另一大特色。除了坚持“三基（基础理论、基本知识和基本技能）和五性（思想性、科学性、先进性、启发性和适用性）”，更注重激发学生的思维，让他们成为自己头脑的主人，批判地看待事物，辩证地对待知识，创造性地预见未来。同时，这版教材也特别注重与五年制教材、研究生教材、专科医师培训教材以及参考书的区别与联系。

以吴阶平、裘法祖、吴孟超、陈灏珠为代表的德高望重的老前辈对第二版教材寄予了殷切期望和悉心指导，教育部、卫生部、国家中医药管理局、国家食品药品监督管理局的各位领导的支持是这版教材不断完善的动力之源。在这里，衷心感谢所有关心这套教材的人们！正是你们的关注，广大师生手中才会捧上这样一本融贯中西、汇纳百家的精品。

八年制医学教材的第一版是我国医学教育史上的重要创举，相信修订后的第二版将不负我国医学教育改革的使命和重任，为培养高层次的具有综合素质和发展潜能的医药卫生人才做出更大的贡献。诚然，修订过程虽然力求完美，但纰漏与瑕疵在所难免，冀望各位领导、同道及师生不吝赐教，以便于这套教材能够与时俱进，不断完善。

是为序。

中国工程院院士
中国医学科学院院长
北京协和医学院院长
刘德培 于庚寅端午佳节
二〇一〇年六月十六日

第2版前言

自2005年由人民卫生出版社出版发行以来，《医学遗传学》一书一再重印，印数不断增加，除作为8年制及7年制临床医学专业的教科书外，其他医学相关专业的师生也都将其作为教材和教学参考用书，在使用过程中深得广大师生的好评，这使我们备受鼓舞，也激励我们以更大的热情投入到第2轮教材的编写工作中。

在第2轮教材编写的过程中，我们广泛采纳广大师生的意见和建议，同时又参照国外同类教材的有益经验，在内容选择、编排体系等方面对原有教材做了较大的改动，删除了免疫遗传和遗传流行病学两个章节，并对余下的14个章节的内容做了不同程度的修改和调整，在每个章节的后面增加了约200个单词的英文小结，力求做到内容更为丰富、翔实，紧跟学科进展，有利于学生知识面的扩展，并熟悉常用的专业英语词汇。

自第1轮教材问世以来，生命科学领域的发展突飞猛进，而作为其核心学科，遗传学的发展更为迅猛，在完成人类基因组全部序列的测定后，以基因组学为龙头，生命科学的各分支学科相互交叉，而生物医学又与物理、化学、数学、纳米、工程学及其他学科相互交叉，相互融合，极大地推动了生命科学的发展。 在医学遗传学领域里，随着功能基因组学的研究进展，新的致病基因不断地被发现，极大地拓展了我们对遗传病本质的认识。 而在表观遗传学的研究方面也不断有新的发现，如非编码小RNA分子在调节真核基因组表达和功能中所起的关键作用等，极大地丰富了我们在遗传学研究方面的知识。 此外，随着医学遗传学研究的不断深入，一些新的研究成果逐渐进入临床应用，我们也不断面临伦理学的问题，如何在遗传学的研究和遗传服务的过程中正确地把握伦理的尺度始终是需要密切注意的问题。 我们在新一轮的教材中对以上问题均有涉及，以使相应的教学能紧跟国际研究的前沿。

参加本轮教材编写的10位教授、专家来自国内8所著名的大学，他们长期工作在医学遗传学教学和科研的第一线，具有很深的学术造诣和丰富的教育经验，在编写过程中大家充分发扬学术民主，各抒己见，对全书的编排和内容的安排提了许多有益的意见和建议，并以认真负责的精神对待教材的编写，使本书能在规定的时间内高质量地完稿，对他们的敬业精神和负责态度表示衷心的感谢。

根据本人的意愿，第1轮教材的副主编傅继梁教授不再参加本轮教材的编写，但他仍十分关心我们的工作，并对第2轮教材，尤其是第十三章表观遗传学的编写工作提出了许多建设性的意见，此外，在上一轮负责免疫遗传和遗传流行病学编写的范丽安教授和沈福民教授也不再参加本轮教材的编写工作了，在此，谨向以上这三位教授表达我们崇高的敬意和衷心的感谢，感谢他们长期以来对医学遗传学教材建设的关心和支持。

因学识与水平的限制，本教材难免有不足之处，希望使用本书的师生们对本书的内容、编排形式等方面多提宝贵意见，使它能更好地推动医学遗传学教学和学科发展。

陈　竺

2010年2月15日

目 录

第一章 遗传学与医学

第一节 健康与疾病的遗传基础

各种生物体包括人体在内，都以其独特的代谢方式利用从周围环境获得的物质，将其改造成为自身可利用的物质，并借以取得能量，维持生命，而将代谢废物排出体外。独特的代谢方式决定于生物体独特的遗传结构。人体独特的遗传结构是进化历程的产物。所谓健康（health），即是受人体遗传结构控制的代谢方式与人体的周围环境保持平衡。遗传结构的缺陷或周围环境的显著改变，都能打破这种平衡，这就意味着疾病（disease，disorder，illness）。在不同疾病的病因中，遗传因素和环境因素所占比重各有不同。外伤、中毒、营养性疾病显然是由环境因素引起，但另有一些疾病则主要是遗传性的，如由突变基因引起的半乳糖血症、苯丙酮尿症等和由染色体畸变引起的 Down 综合征、Turner 综合征等。这些疾病只发生于有异常基因或有异常染色体数目或结构的个体。还有一些异常遗传结构虽然改变了个体的代谢，但在一般生活条件下仍可为个体所耐受，只在接触特殊环境条件时才发病，如葡萄糖-6-磷酸脱氢酶（G-6-PD）缺乏者在食用蚕豆或服用伯氨喹等药物后发生溶血危象。许多常见病如高血压、糖尿病、肿瘤、风湿病、消化性溃疡等介于两者之间。这些疾病有一定的遗传因素，家族发病率高于人群发病率，但其发病都以一定的环境条件为其诱因，遗传因素在其中所起作用程度各异。这些疾病的遗传因素是若干基因微小作用的累加效应。

因遗传因素而罹患的疾病称为遗传性疾病或简称遗传病（inherited disease，genetic disorder）。遗传因素可以是生殖细胞或受精卵内遗传物质的结构和功能的改变，也可以是体细胞内遗传物质结构和功能的改变。大多数遗传病为先天性疾病（congenital disease）。所谓先天性疾病是指婴儿出生时即显示症状，如尿黑酸尿症、血友病、Down 综合征等。但先天性疾病不一定是遗传病，如胎儿在宫内感染天花造成出生时脸上有瘢痕，母亲怀孕早期感染风疹病毒致使胎儿患有先天性心脏病，孕妇服用

反应停(thalidomide)引起胎儿先天畸形。同样,有不少遗传病出生时毫无症状,要到一定年龄才发病,如肌营养不良症到儿童期发病,Huntington 舞蹈病一般发病于 25~45 岁,痛风好发于 30~35 岁。遗传病往往表现为家族性疾病(familial disease),在亲代和子代中均有患者,或在正常父母所生同胞中出现一个以上的患者。遗传病也可能呈散发性,这是正常亲代的生殖细胞发生基因突变或染色体畸变而使子代得病。有些遗传病还可能有不外显的亲代,患者的出现也可呈散发性。但同样,家族性疾病并不一定就是遗传病。一个家族有多个成员患同一疾病(如结核病、肝炎)可能系由共同的生活环境所引起。夜盲也常有家族性,但它显然并非遗传病,而仅为维生素 A 缺乏所致。

据 WHO(1973)技术报告,在加拿大 Montreal 儿童医院 1969~1970 年的 1146 名住院患儿中,与遗传有关的疾病占 29.4%。新中国成立 50 余年来,随着儿童保健事业的发展,婴儿死亡率大为降低。在婴儿死亡原因中,与遗传有关的疾病所占比例越来越大。如以北京市为例,先天性心脏病在 1951 年全部婴儿死亡原因中仅占 2.6%,1974~1976 年则上升到 30%,占各种疾病死亡中的第二位。与此相似,先天畸形在北京市 1974~1976 年 15 岁以下儿童五种主要疾病死亡原因中已占首位。与遗传有关的先天畸形,包括先天性心脏病、大脑发育不全、消化道畸形、脊柱裂、无脑儿、脑积水、多发畸形等,约占先天畸形死亡总数的 90%。恶性肿瘤包括白血病、脑瘤和神经母细胞瘤,约占恶性肿瘤死亡总数的 70%。先天畸形和恶性肿瘤两者合计要占到儿童死因的 30% 以上。可见遗传病和先天畸形已成为儿童死亡的主要原因。另一方面,随着人们平均寿命的延长,人群中老年人所占的比例上升,由遗传和环境因素综合作用引起的高血压、心血管疾病、关节炎、恶性肿瘤等老年性疾病的比重也在逐年增加。无疑,随着传染病得到控制,我国人群的疾病谱(disease spectrum)已经改变,在各个年龄组中,遗传病的重要性越来越显著。

各种遗传病的发病率在不同人群中是有差别的。如在地中海沿岸的意大利人和希腊人中,β 地中海贫血、地中海型 G-6-PD 缺乏症和家族性地中海热的发病率较高。在我国和其他东方人中,则是 α 地中海贫血、中国型 G-6-PD 缺乏症和成年型乳糖酶缺乏的发病率较高。据 Scriver 等(1978)综合各家报道,在白人各种疾病患者中,遗传病发生率总的估计值是:①染色体畸变见于自发流产儿约 50%,见于死产儿约 5%,见于 7 年内死亡儿童约 5%,见于成活婴儿约 0.5%;②儿科住院患儿约 12% 患有染色体病、单基因病或多基因病;另外 18% 患有多基因先天畸形;③成人住院患者至少 12% 患有显著遗传成分的疾病;④在严重智力落后者中约 15% 患有单基因病,约 45% 有遗传成分。

第二节 医学遗传学发展简史

一、遗传病的早期认识

关于遗传的概念至少可追溯到古希腊 Hippocrates 时代之前,当时人们就已经认识到某些疾病可能在家庭中传递。大约 1500 年之前,犹太教法典就有对“易出血者”的某些男性家属免除割礼的规定,证明人们已经认识了血友病的遗传规律。18 世纪 Maupertuis 研究了多指(趾)及皮肤和毛发缺乏色素者(白化症)的家系,指出这两种症状有各自不同的遗传方式。

1859 年 Boedeker 首先确诊尿黑酸尿症,这是最早报道的先天性代谢病。

现代遗传学的奠基人是 19 世纪奥地利的僧侣 Mendel(1822~1884)。1866 年他发表了豌豆杂交实验,发现黄豌豆植株与绿豌豆植株杂交,子代都是黄豌豆,黄对绿是显性。子代自花授粉时,子代豌豆有黄有绿。Mendel 根据实验结果认为,遗传性状是由成对的遗传因子决定的。在生殖细胞形成时,成对的遗传因子要分开,分别进入两个生殖细胞中去。这被后人称为 Mendel 第一定律或分离率(law of segregation)。Mendel 同时认为,在生殖细胞形成时,不同对的遗传因子可以自由组合,这就是 Mendel 第二定律或自由组合律(law of independent assortment)。这两个定律是 Mendel 遗传因子学说的中心内容。此后,无数实验都证明了此学说的正确性。但是,Mendel 的工作直到 1900 年才被发现。随即有人试图把 Mendel 遗传因子学说应用于人类。Farabee(1903)指出短指(趾)为显性性状,

这是人类显性遗传的第一例。1901 年 Garrod 描述 4 个尿黑酸尿症家系，共有 11 个患者，其中至少有 3 个患者的父母为表亲，他们看来都是正常的。遗传学家 Bateson 提示 Garrod，尿黑酸尿症属隐性性状。Bateson 认为，表亲由于有共同的外祖父母，他们更有可能具有同样的遗传因子，因此在具有两个隐性因子的患者的父母中，预期近亲婚配的频率较高。Garrod 的尿黑酸尿症正是如此。此后发现人体许多遗传性状都符合于 Mendel 遗传，但当时不少学者试图把各自的观察结果套到 Mendel 的显隐概念中去，这就显然把事情过于简单化了，以至不得不提出各种补充假设。例如 Davenport 坚信智力缺陷一般为隐性遗传性状。这显然是片面的，因为智力缺陷有多种类型，其中有些是按 Mendel 性状遗传的，但大多数属多因子效应。实际上，对大多数疾病来说，遗传因素和环境因素都有作用。

1903 年 Sutton 和 Boveri 分别注意到 Mendel 遗传因子的行为跟生殖细胞形成和受精过程中染色体的行为完全平行，于是两人分别提出，遗传因子就在染色体上，这就是染色体遗传学说。1909 年 Johannsen 将遗传因子改称为基因(gene)。

在 1905 年以前，大多数遗传学实验是在植物身上进行的。1905 年 Castle 用果蝇(Drosophila)进行了遗传学实验。果蝇容易饲养，一年可繁殖 20～25 代。此外，果蝇只有 4 对染色体，研究方便。1910 年左右，美国哥伦比亚大学的 Morgan 和他的学生 Sturtevant、Bridges 和 Muller 开始研究果蝇性状的遗传方式，发现果蝇的性状可分为 4 个遗传连锁群(linkage group)，这恰和果蝇的染色体对数相一致。染色体是遗传的传递单位，因此一条染色体上的连锁基因也是一起传递给子代的，这就是连锁律(law of linkage)。但是连锁不是绝对的。在生殖细胞形成过程中，同源染色体之间有时可交换一个片段，使连锁基因发生重新组合，这就是交换律(law of crossing-over)。

Johannsen 不仅首先将遗传因子改称为基因，而且还是区别基因型(genotype)和表现型(phenotype)的创始人。基因型指个体的遗传结构；表现型系指环境条件与基因型相互作用而使该个体呈现的性状。不过对于人体，早在 1875 年，Galton 就已区分了先天(nature)与后天环境(nurture)的影响。Galton 认为，由于一卵双生而有相同的遗传结构即有同样的基因型，但在不同的环境中生长可有不同的表现型。他对体质和才能的遗传特别感兴趣。他把回归系数这一统计概念引进遗传学，借以估计各种亲属间的相似程度。Galton 的工作为以后人类遗传学中涉及数学问题的研究奠定了基石。Galton 还提出改进人和动物遗传素质的概念，并称之为优生学(eugenics)，由此导致 20 世纪 30～40 年代反科学、反人道的优生运动。

Adams 于 1914 年发表论文，讨论到先天性疾病、家族性疾病和遗传性疾病的差别，遗传病与发病年龄、环境因素、近亲结婚之间的关系等，其中涉及了有关遗传病的一些基本问题，并逻辑地推论到遗传的一些基本原理。

二、医学遗传学的兴起

医学遗传学(medical genetics)是临床医学与遗传学相互渗透的一门边缘科学，是人类遗传学的一个组成部分。医学遗传学揭示了人类纷繁的变异库，为人类遗传学研究提供了丰富的素材。20 世纪 50 年代以来，医学遗传学有了迅猛的发展，这主要是由于生物化学、细胞遗传学、免疫学与分子遗传学实验技术的发展起了推动作用。

(一) 医学遗传学与生物化学实验技术的发展

前已述及，Garrod 研究尿黑酸尿症并推测患者体内的尿黑酸是酪氨酸的降解产物，由于先天性酶缺乏而不能被进一步氧化，因而在体内累积并随尿排出。此后，La Du 等(1958)发现患者活检肝组织中缺乏尿黑酸氧化酶，从而证实了 Garrod 的假设。现已检出许多由各种遗传性酶缺乏引起的代谢病。20 世纪 50 年代以来，生化实验技术和分析方法的发展，提高了对先天性代谢病的研究和临床诊断的水平。例如由层析法检出尿液中的异常代谢产物，由电泳技术检出异常血红蛋白分子，淀粉凝胶电泳可检出包括酶在内的蛋白质的结构异常。这就使医学遗传学在理论研究和实际应用两方面都向前跨进了一大步。

在理论研究上，最引人注目的进展是对血红蛋白的研究。Pauling 等(1949)在研究镰形红细胞贫血时发现患者有一种异常血红蛋白分子 HbS，其电泳性质不同于正常的 HbA，从而提出了分子病(molecular disease)的概念。1954 年 Ingram 创立"指纹法"，查明 HbSβ 链第 6 位氨基酸是缬氨酸而不是正常的谷氨酸，由此展开了对血红蛋白分子病的深入研究。

在实际应用上，开辟了治疗某些遗传病的有效途径。苯丙酮尿症的治疗标志着这方面的重大进展。1953 年 Bickel 等提出，通过控制新生儿的苯丙氨酸摄入量，有效地防止苯丙酮尿症的发展，并取得治疗效果。此项工作对开展早期检出遗传病的研究以及寻找防治和控制先天性代谢病的有效方法起了推动作用。

20 世纪 50 年代中期，发现乙酰胆碱敏感是由于血清胆碱酯酶缺乏所致，伯氨喹引起药物性溶血是由于 G-6-PD 缺乏所致。这些发现说明，药物反应有受遗传控制的代谢基础。1959 年 Vogel 提出药物遗传学(pharmacogenetics)一词。进入 20 世纪 70 年代后，这一概念进一步扩展，不仅药物反应要考虑遗传基础，而且对一切环境因子的反应，包括食物反应在内，也必须考虑遗传基础。1971 年 Brewer 提出生态遗传学(ecogenetics)这一术语。对食物、药物和毒物的反应都必须从遗传基础加以认识。

（二） 医学遗传学与细胞遗传学染色体实验技术的发展

1956 年 J. H. Tjio(蒋有兴)和 Levan 观察人胎肺组织培养细胞，首先正确地鉴定人体体细胞的染色体数目为 46 条。同年 Ford 和 Hamerton 观察到人的精原细胞有丝分裂中期的染色体，从而证实了上述结果。但是染色体分析之所以能够普遍开展，还要归功于细胞培养和制片技术上的一系列突破。1952 年 T. C. Hsu(徐道觉)发现，分裂细胞经低渗处理，可使细胞膨胀，染色体分散，便于观察。1956 年 J. H. Tjio(蒋有兴)利用秋水仙碱(colchicine)阻止细胞进入分裂后期，使分裂中期图形增多。1960 年 Nowell 等应用植物血凝素(phytohaemagglutinin，PHA)使体外培养的人体淋巴细胞母细胞化而进入分裂。同年，Moorhead 等综合应用各项新技术，建立人体外周血体外培养和染色体制片等一整套实验技术，从而有了使染色体研究简便可靠的方法。

在人体染色体数目得到正确鉴定之后，染色体分析技术即被迅速应用于临床。1959 年就有三个发现：Lejeune 等发现 Down 综合征患者有 47 条染色体，即多 1 条小型近端着丝粒染色体(第 21 号染色体)；Ford 发现 Turner 综合征妇女只有 1 条 X 染色体；Jacobs 和 Strong 发现 Klinefelter 综合征男子的性染色体是 XXY。1960 年美国费城(Philadelphia)研究小组在慢性粒细胞性白血病(chronic myelocytic leukemia，CML)患者的细胞里第一次发现了特定的染色体结构畸变，称之为费城染色体或 Ph^1 染色体。随后又发现了其他许多染色体综合征和一些肿瘤的标记染色体。20 世纪 60 年代末 Caspersson 等发现，植物染色体在用荧光染料氮芥喹吖因(quinacrine mustard)染色时，染色体的不同区域显示强弱不等荧光。1970 年再将此技术应用于人体染色体，揭示了人体各条染色体独特的荧光带型。由此提高了染色体分析的精确性，并发现了不少新的染色体综合征。

（三） 医学遗传学与免疫学实验技术的发展

免疫学实验技术的发展扩大了遗传病的概念，并为疾病防治带来了新的方法。1900 年 Landsteiner 发现了 ABO 血型。在这以后的半个世纪中，利用红细胞凝集试验先后发现了十几个血型系统，为临床输血配型奠定了基础。1941 年 Levine 等提出，胎儿红细胞增多症(新生儿溶血症)系由胎母红细胞抗原不相容引起同种免疫(alloimmunization)所致。1952 年，Dausset 和 Nenna 在多次输血患者的血液中发现白细胞凝集素，1958 年检出第一个白细胞抗原 Mac(HLA-A2 + A28)。1964 年 Terasaki 和 McClelland 设计微量淋巴细胞毒试验以取代白细胞凝集试验。此后，免疫遗传学研究揭示了人体高度多态性的 HLA 系统，使器官移植供、受体配型有了可能。1950 年 Glanzmann 和 Rinicker 描述了严重联合免疫缺乏综合征。1952 年 Bruton 报道了单纯为体液免疫缺乏的低丙种球蛋白血症。现已揭示了一系列遗传方式各异、临床表现多种多样的遗传性免疫缺乏病。

（四） 医学遗传学与分子遗传学 DNA 实验技术的发展

20 世纪 70 年代初，随限制性内切酶的发现及 DNA 分子杂交技术的建立，分子遗传学进入基因

工程阶段，并为解决临床问题提供了新的手段。Y. W. Kan（简悦威）等（1976）、Wong 等（1978）及 Dozy 等（1979）应用 DNA 实验技术，就胎儿羊水细胞 DNA 作出 α 地中海贫血出生前诊断。由于限制性内切酶在消化 DNA 时其切割部位的核苷酸顺序有严格的特异性，因此在突变导致 DNA 的核苷酸顺序发生改变时，原有的内切酶切割部位可能消失，也可能出现新的切割部位。Y. W. Kan（简悦威）等（1978）据此就胎儿羊水细胞 DNA 作出镰形红细胞贫血症出生前诊断。近年来，苯丙酮尿症、血友病等都已能在 DNA 水平上作出诊断。

综观 20 世纪 50 年代以来医学遗传学的发展，人们研究与疾病发生有关的各种生物学变异，从表现型变异，蛋白质变异，进行到 DNA 变异。历届国际人类遗传学大会清楚地反映了研究水平的不断深入。20 世纪 70 年代中期兴起分子遗传学，极大地促进了医学遗传学的发展，揭示癌基因（oncogene）和（或）肿瘤抑制基因（tumor suppressor gene）的突变是肿瘤发生的分子基础，从而确定肿瘤是一种体细胞遗传病。体细胞突变也可能是自身免疫性疾病和衰老过程的分子基础。

随着分子遗传学的发展，20 世纪 90 年代初，基因治疗（gene therapy）进入了临床试验阶段。所谓基因治疗就是将某个正常基因导入患者体内细胞中使之表达，对患者缺乏的或异常的某种蛋白质提供其正常表达产物，从而起到治疗作用。由腺苷脱氨酶（adenosine deaminase，ADA）缺乏引起的严重联合免疫缺乏症（severe combined immunodeficinecy，SCID）和由凝血Ⅸ因子缺乏引起的 B 型血友病，基因治疗的临床试验都已得到令人鼓舞的治疗效果。

分子遗传学的发展导致了反求遗传学（reverse genetics）的新趋势。这就是在不知道某种遗传病蛋白质异常的情况下，直接寻找致病的 DNA 变异，进而揭示这种 DNA 变异所导致的蛋白质异常。这就使遗传学研究从表现型到基因型这条经典路线转变成为从基因型到表现型的反求路线。近年来，在反求遗传学思路指导下，遗传学家们对一些不明原因的遗传病进行了 DNA 标记连锁分析，使 DNA 标记逐渐逼近，最终找到致病基因。遗传学家们由此找到了 Duchenne 肌营养不良（Duchenne muscular dystrophy，DMD）的肌营养不良蛋白（dystrophin）基因和囊性纤维变性（cystic fibrosis，CF）的跨膜调节蛋白（transmembrane regulator）基因等。

遗传病的分子遗传学研究，正在使医学遗传学走向 21 世纪的大发展。1986 年，诺贝尔奖金获得者 Dulbecco 提出，如果我们希望对肿瘤有更多的了解，这就必须集中于细胞的基因组，最有用的是从细胞的基因组测序开始。经过学术界几年的争论，1990 年美国国会批准十五年（1991～2005）拨款 30 亿美元的人类基因组计划（human genome project，HGP）。计划通过三部曲，即连锁图（遗传图）、物理图和基因组测序，揭示人类基因组 DNA 30 亿碱基对的全序列。HGP 是生物医学领域的阿波罗登月计划，它将给 21 世纪的生物医学科学带来一场遗传学革命。由于 HGP 意义重大而影响深远，它引起各国政府高度重视，纷纷投入大量资金推进 HGP 研究，使其研究进展一再超前。2000 年 6 月 26 日美国总统克林顿和英国首相布莱尔宣布人类基因组序列工作草图诞生。2001 年 2 月 15 日美、英、日、法、德、中六国国际人类基因组测序联合体发表了根据人类基因组 94% 序列草图作出的初步分析。2004 年 10 月 21 日，《Nature》杂志公布了人类基因组的完成序列。这个人类分子遗传学的登峰之作（表 1-1）必将引导 21 世纪的生物医学科学结出丰硕成果，进一步造福于人类。

表 1-1　医学分子遗传学大事记

年份	重大事件	主要学者
1869	首次分离 DNA	Miescher
1944	证明在细菌转化过程中携带遗传信息是 DNA 而不是蛋白质	Avery
1953	揭示 DNA 双螺旋结构	Watson，Crick
1961	发现 DNA 复性，从而建立了核酸杂交反应的特异性和可行性	Marmur，Doty

续表

年份	重大事件	主要学者
1962	首次证明存在 DNA 限制性内切酶	Arber
1966	阐明 DNA 遗传密码	Nirenberg, Ochoa, Khorana
1967	发现 DNA 连接酶	Gellert
1970	首例试管内合成基因	Khorana
1972 ~ 1973	DNA 克隆技术	Boyer, Cohen, Berg
1975	凝胶转移杂交以检测特异性 DNA 序列	Southern
1975 ~ 1977	快速 DNA 测序技术	Sanger 和 Barrell, Maxam 和 Gilbert
1977	首例人类基因克隆	Shine
1978	首例 RFLP	Kan(简悦威)
1978	首例 DNA 诊断	Kan(简悦威)
1981	人类线粒体 DNA 测序	Anderson
1985	首例 DNA 指纹	Jeffreys
1985	发明聚合酶链反应	Mullis, Saiki, Erlich
1986	提出肿瘤问题有待人类基因组测序	Dulbecco
1990	首例腺苷脱氨酶缺乏症临床基因治疗	Anderson
1991	人类基因组研究 15 年规划启动	Waston, Collins
1994	人类基因组连锁图	Murray, Weissenbach, White, Ward, Dausset
1998	人类基因组物理图	Deloukas, Schuler, Gyapay, Beasley
2001.2.15	根据人类基因组 94% 序列草图作出初步分析	美、英、日、法、德、中六国国际人类基因组测序联合体, Celara 公司
2004.10.21	人类基因组完成序列,该序列覆盖了约 99% 的常染色质区域,错误率小于 1/10 万	美、英、日、法、德、中六国国际人类基因组测序联合体

医学遗传学发展至今,已经成为一门涉及基础与临床的综合性学科,其分支学科见表 1-2。

表 1-2　医学遗传学的分支学科

细胞遗传学(cytogenetics)	体细胞遗传学(somatic cell genetics)
生化遗传学(biochemical genetics)	癌肿遗传学(cancer genetics)
分子遗传学(molecular genetics)	群体遗传学(population genetics)
药物遗传学(pharmacogenetics)	遗传流行病学(genetic epidemiology)
免疫遗传学(immunogenetics)	临床遗传学(clinical genetics)
行为遗传学(behavioral genetics)	基因组学(genomics)
生态遗传学(ecogenetics)	药物基因组学(pharmacogenomics)
辐射遗传学(radiation genetics)	

第三节　遗传性疾病的分类

遗传病可分为五大类。在分析一种疾病的遗传基础时,首先要确定它属于这五大类中的哪一类。

一、染色体病

人类正常体细胞具有二倍体数46条染色体。如果在生殖细胞发生和受精卵早期发育过程中发生了差错，就会产生整条染色体或染色体节段超过或少于二倍体数的个体，表现为种种先天发育异常。Down综合征即由于第21号染色体多了1条，成为21三体性。染色体病（chromosome disorders）通常不在家系中传递，但也有可传递的。已知的染色体病有300多种。出生时染色体病发生率约为7‰。在妊娠前3个月的自发性流产中，染色体畸变大约要占到一半。

二、单基因病

单基因病（single-gene disorders）起因于突变基因。在一对同源染色体上，可能其中一条带有突变基因，也可能两条染色体对应位点都是突变基因。单基因病通常呈现特征性的家系传递格局。个别单基因病均属罕见，其发生率的上限约为2‰。但发现的病种越来越多，从这一类疾病来说并非罕见（表1-3）。

表1-3　在线人类孟德尔遗传数据库统计资料（2010年3月4日）

类　型	常染色体	X连锁	Y连锁	线粒体	总计
已知序列的基因数	1 2338	609	48	35	1 3030
已知序列和表型的基因数	323	19	0	2	344
已知分子基础并有表型描述的基因数	2453	216	4	26	2699
分子基础不明，但符合孟德尔表型或位点的基因数	1644	142	5	0	1791
其他，类似孟德尔表型的基因数	1865	136	2	0	2003
总计	1 8623	1122	59	63	1 9867

三、多基因病

多基因病（polygenic disorders）亦称复杂疾病（complex disease），起因于遗传素质和环境因素，包括一些先天性发育异常和一些常见病。多基因病有家族聚集现象，但无单基因病那样明确的家系传递格局。

四、线粒体基因病

线粒体DNA为呼吸链部分肽链及线粒体蛋白质合成系统rRNA和tRNA编码。这些线粒体基因突变可致线粒体基因病（mitochondrial genetic disorders），随同线粒体传递，呈细胞质遗传。

五、体细胞遗传病

已知肿瘤起因于遗传物质的突变。癌家族可有家族性肿瘤遗传易感性，但体细胞癌肿病灶具有克隆性（clonality），其形成必以体细胞遗传物质突变为直接原因，故肿瘤属于体细胞遗传病（somatic cell genetic disorders）。有些先天畸形亦属此类。

一种疾病是否有遗传病因，有时不易确定。即使感染性疾病也还有一个遗传素质问题。下列事实提示存在某种遗传病因：

1. 患者有特征性表现型（通常伴有智力障碍）和染色体异常，有或无同一疾病或有关疾病的家族史。
2. 在排除环境因素的前提下，亲属中有一定比例的患病者。
3. 在非血缘成员（配偶）中不出现患者。

4. 在未知有触发因素的情况下，患者有特征性的发病年龄和病程。
5. 同卵双生的同病率高于异卵双生。

第四节　医学遗传学与遗传医学的任务和展望

医学遗传学的任务在于揭示各种遗传性疾病的遗传规律、发病机理、诊断和防治措施。遗传医学(genetic medicine)则为遗传病患者提供临床服务，包括遗传病的诊断、治疗、筛查、预防、咨询、随访等。其最终目的在于尽可能减少遗传病患者的痛楚，使他们尽可能享有平安的幸福人生。改革开放以来，我国大城市的医院结合计划生育逐步建立起婚前检查门诊和遗传咨询门诊。临床各科的遗传医学服务也日益受到重视。遗传医学必将在我国医学现代化的过程中作出应有的贡献。

根据国际人类基因组测序联合体对人类基因组 DNA 完成序列的分析，人类基因组只有 2～2.5 万个编码蛋白质的基因，仅占人类基因组全序列的 1.1%～1.4%，发现人类基因组有着 1.42×10^6 单核苷酸多态性(single nucleotide polymorphism，SNP)。在人类基因组测序过程中随着多态性序列标志发现得越来越多，定位克隆(positional cloning)寻找疾病基因的进度日益加快(表 1-4)。所有单基因病的致病基因必将全部得到鉴定。21 世纪医学遗传学研究的重点将是多基因复杂病和癌肿，它们涉及多个遗传基因和环境因素之间的相互作用。1.42×10^6SNP 是研究复杂病和肿瘤的极好的多态性遗传标志。多基因复杂病和肿瘤无疑也涉及基因组的调控机制。这也是 21 世纪医学遗传学必须研究的课题。

表 1-4　应用人类基因组序列草图定位克隆的疾病基因*

基因位点	遗传病	主要学者
BRCA2	乳房癌易感性	Wooster
AIRE	自身免疫性多腺体综合征 1 型	Nagamine
PEX1	过氧化物酶体生物发生病	Reuber，Portsteffen
PDS	Pendred 综合征	Everett
XLP	X 连锁淋巴细胞增生病	Coffey
DFNA5	非综合征耳聋	Van Laer
ATP2A2	Darier 病	Sakuntabhai
SEDL	迟发型 X 连锁脊椎骺发育不良	Gedeon
WISP3	进行性假风湿样发育不良	Hurvitz
CCM1	大脑空洞畸形	Sahoo
COL11A2/DFNA13	非综合征耳聋	McGuirt
LGMD2G	肢带肌营养不良 2G 型	Moreira
EVC	Ellis-Van Creveld 综合征 Weyer 肢端骨发育障碍	Ruiz-Perez
ACTN4	家族性病灶节段性肾小球硬化症	Kaplan
SCN1A	全身性癫痫伴发热发作 2 型	Escayg
AASS	家族性高赖氨酸血	Sacksteder
NDRG1	遗传性运动和感觉神经病	Kalaydjieva-Lom
CNGB3	全色盲	Kohl，Sundin
MUL	Mulibrey 侏儒症	Avela
USH1C	Usher 综合征 1C 型	Kerpy，Bitner-Glindzicy

续表

基因位点	遗 传 病	主 要 学 者
MYH9	May-Hegglin 异常	Kelley, May-Hegglin
	Fetchner 综合征	Fetchner 综合征联合体
	非综合征遗传性耳聋 DFNA17	Lalwani
PRKAR1A	Carney 复合征	Kirschner
SCA10	脊髓小脑性共济失调 10 型	Matsuura
OPA1	视神经萎缩	Delettre
XLCSNB	X 连锁先天性静止型夜盲	Purch
FGF23	低磷酸盐血症	常染色体显性低磷酸盐血症联合体
GAN	巨轴索神经病	Bomont
AAAS	三 A 综合征	Tullio-Pelet
HSPG2	Schwartz-Jampel 综合征	Nicole

* 引自 International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature, 2001, 409: 860-921

在功能基因组时代,医学需要了解遗传、表观遗传的因素以及环境因素在生理和病理过程中的作用,并对其总体信息进行表征。蛋白编码基因序列的揭示,为利用转基因动、植物生物学反应器制备相应蛋白质提供了条件。结构基因组学(structural genomics)自然已向功能基因组学(functional genomcis)和蛋白质组学(proteomics)发展。人类基因组 DNA 序列绝大多数均为非编码序列。这些非编码序列究竟有何生物学意义,也是 21 世纪人类与医学遗传学必须面对的问题。

人类基因组这本生命天书的揭露,必将使 21 世纪的医学发生革命性的变化。新世纪的医学将是循证的、个体化的系统医学,系统生物学将成为新世纪医学发展的核心驱动力。对人体健康和病理状态的充分了解,需要对人体的系统结构和动力学进行深入的研究;对重要人类疾病的控制和预防,需要发展新的、系统的模式,包括从机理性研究到临床诊断、治疗。基因芯片(chip)亦称微阵(microarray)进入临床服务时,不仅可以高效进行分子诊断,而且可以鉴定每个人基因组的表达格局,即基因组的生物学密码。临床医师可以根据每个人的生物学密码,制定个人特异的治疗方案,避免药物的毒副反应。临床医师还可以根据每个人的生物学密码判断多基因复杂病的发病风险,通过改进生活方式,防止发病,使医疗服务从治病走向防病。随着更多位点的致病基因和易感基因被揭示,体细胞基因治疗将成为临床的常规疗法。我们还应该充分重视公共卫生事业。因为医学强调的是诊断和治疗,日益强调个体化,如果有问题,影响的只是个人,少部分人;而公共卫生强调的是预防,一旦出问题,就会影响到一大群人,甚至整个社会。

高科技的基因组医学在为患者做好事的同时,无疑也存在个人遗传信息泄密或被误用的风险。因此,临床医师必须注意为患者保守机密,维护患者知情同意的权益,反对基因增强(gene enhancement)和胎儿选择以防止滑向反科学和反人道的新优生运动。

医学正在进行一场影响深远的遗传学革命。临床医师必须掌握发展中的医学遗传学和医学伦理学,才能为患者、为社会真正做好事。

Summary

The health of human being depends on a balance of the genetically regulated metabolism and the changing environmental conditions.Defects in genetic information embedded in the deoxy-

ribonucleic acid(DNA) sequence of the genome and/or profound changes of the environment disrupt this balance and lead to diseases.Analysis of the human disease spectrum reveals an interesting phenomenon: the weight of the genetic and environmental factors varies significantly according to the etiology of different diseases.

Medical genetics is a relatively new discipline as a result of mutual infiltration of clinical medicine and genetics.It is an integrated part of human genetics.

Genetic disorders are generally divided into five categories: chromosomal disorder、single gene disorder、polygenic disorder、mitochrondral disorder and somatic cell genetic disorder.

Medicine is undertaking a revolution led by genetics.Clinical doctors can contribute to the patients and society only if they can master medical genetics and medical ethics.

(陈 竺)

参考文献

1. 谈家桢. 医学遗传学//中国医学百科全书编辑委员会. 中国医学百科全书医学遗传学. 上海:上海科学技术出版社, 1984,1-3
2. Edwards JH. The importance of genetic diseases and the need for prevention. Phil Trans R Soc Lond B,1988,319:211-217
3. McKusick VA. Foreword//Gelehrter TD, Collins FS. Principles of Medical Genetics. Baltimore: Williams and Wilkins, 1990, 5-6
4. McKusick VA. Human genetics: The last 35 years, the present, and the future. Am J Hum Genet, 1992, 50:663-670
5. Beaudet AL. Making genomic medicine a reality. Am J Hum Genet, 1999, 64:1-13
6. Collins FS. Medical and societal consequences of the human genome project. N Engl J Med, 1999, 341(1):28-37
7. van Ommen GJB, Bakker E, den Dunnen JT. The human genome project and the future of diagnostics, treatment, and prevention. Lancet, 1999, 354(suppl 1):5-10
8. Worton RG. On discovery, genomes, the society, and society. Am J Hum Genet, 2001, 68(4):819-825
9. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature, 2001, 409:860-921

第二章　DNA 与遗传信息传递

DNA(deoxyribonucleic acid)即脱氧核糖核酸，作为遗传物质，储存着大量遗传信息，精确地编码细胞生长、分裂、分化和对内外环境反应的所有指令。认识 DNA 的结构及其特征是医学遗传学和分子医学的基础。本章主要介绍 DNA 结构、功能及其传递。

第一节　DNA 结构与特征

一、DNA 分子结构

在 20 世纪早期，虽然已经知道染色体决定遗传性状，其成分有 DNA 和蛋白质，但倾向于染色体的蛋白质为遗传物质。1944 年，Avery、Macleod 等科学家发现，当毒性光滑型肺炎球菌死株和无毒性粗糙型肺炎球菌活株混合后，部分粗糙型菌株转化为光滑型，证明转化实验中的"转化因子"不是蛋白质，而是 DNA，因此 DNA 是遗传物质。

1953 年，Watson 和 Crick 应用精细的 X 射线晶体衍射数据，推测出 DNA 分子的双螺旋结构。如图 2-1 所示，DNA 的基本成分为核苷酸(nucleotide)，即一个戊糖、一个碱基和三个磷酸构成一个核苷酸。磷酸连接了一个糖的 3′碳原子和相邻的另一个糖的 5′碳原子，通过这样 3′,5′-磷酸二酯键(3′,5′-phosphodiester bond)将核苷酸连成多核苷酸链；每条多核苷酸链的一个末端为 5′末端(5′ end)，另一个末端为 3′末端(3′ end)。两条多核苷酸链相互平行反向缠绕构成 DNA 双螺旋大分子，一条链的方向为 5′→3′，另一条链的方向相反为 3′→5′。碱基有嘌呤和嘧啶两种，嘌呤包括腺嘌呤(adenine,

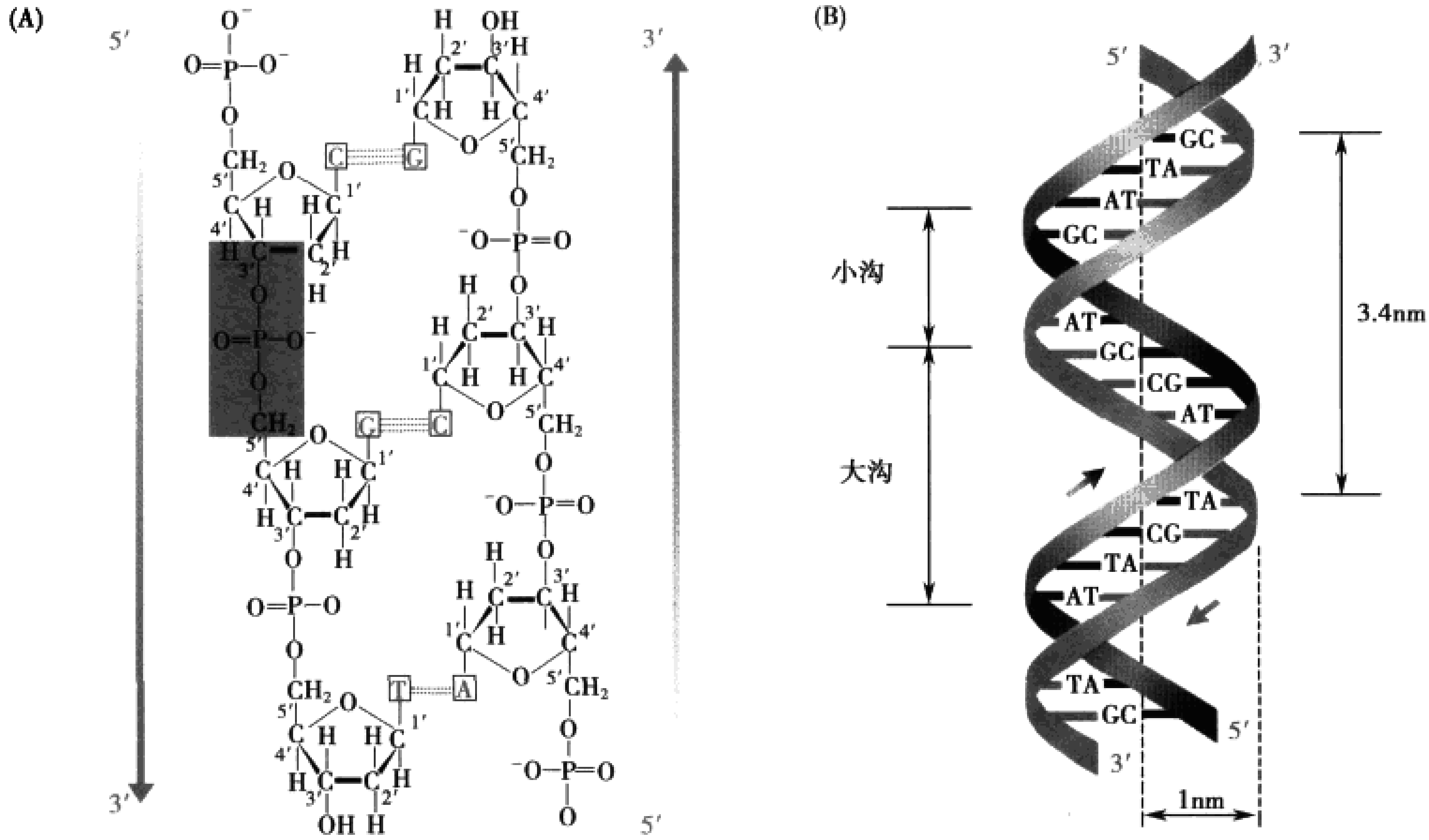


图 2-1 DNA 的双螺旋结构

A)和鸟嘌呤(guanine,G);嘧啶包括胞嘧啶(cytosine,C)和胸腺嘧啶(thymine,T)。两条多核苷酸链通过 A 和 T,G 和 C 之间的氢键相结合,使 DNA 双链彼此互补;A 与 T 间为 2 个氢键,G 与 C 间为 3 个氢键,每一个称为碱基对(base pair,bp),这种 A 与 T、G 与 C 的配对原则说明了特定物种 DNA 中,A 和 T 的百分率相同,G 和 C 的百分率相同;因此,可以通过碱基中 GC 的百分率推算出碱基的组分。例如,某种细胞 DNA 的 GC 含量占 42%,其碱基组分是:G,21%;C,21%;A,29%;T,29%。利用两条 DNA 链的互补(complementary),可以从一条链的组成而获得另一条链的组成。如果一条链为 5′-CAGTTCA-3′,则另一条链为 5′-TGAACTG-3′。

遗传信息储存在 DNA 链中碱基的线性序列(一级结构,primary structure)上。通常按照 5′→3′方向写出一条链的碱基序列,此方向是 DNA 复制新 DNA 分子的方向。当描述一条 DNA 链上两个相邻碱基时,通常插入一个"p"来表示一个连接的磷酸二酯键,例如 CpG 表示在同一条 DNA 链上一个胞苷连接到一个相邻的鸟苷上;而一个 CG 碱基对表示一条 DNA 链上的一个胞嘧啶与互补链上的一个鸟嘌呤形成氢键。

DNA 分子结构表明了如下的生物学意义:

1. DNA 分子的碱基序列储存了大量遗传信息　若一个 DNA 分子长度为 n 个碱基,就可能有 4^n 种排列序列。在 DNA 链上,三个相邻碱基构成了遗传密码的单位,4 种碱基以这种形式组成 $4^3=64$ 个遗传密码,人类的全部遗传信息就是以碱基的不同排列次序蕴藏在全部 DNA 序列之中。

2. DNA 分子的双螺旋碱基互补结构是 DNA 复制和修复的基础　DNA 双链中的每条链都可作为合成一条新链的模板;因此子代 DNA 双螺旋包含一条亲链和一条合成的新链,故称为半保留复制(semiconservative replication)。细胞中 DNA 分子受到损伤时,被破坏或缺损的碱基,可在 DNA 修复酶的作用下,以互补链为模板按碱基互补原则进行修复而替代受损的碱基。

3. DNA 分子的双链互补性是现代分子生物学核心技术——"分子杂交"的基础　正因为 DNA 的碱基互补这一特性,使单链 DNA 可以在复杂的 DNA 混合物中找到与其互补的 DNA 序列。Southern 印迹,Northern 印迹,PCR 技术,人工合成 DNA,以至现今的 DNA 芯片技术等,都是依据碱基互补配对原理,实现分子识别,产生出各种检测和分析基因功能的研究方法。

4. 双螺旋结构形成的大沟是 DNA 与蛋白质相互作用的结构基础　两条多核苷酸链相互缠绕的双螺旋分子包含大沟和小沟结构，基因转录时 DNA 与各种转录因子的相互作用就是转录因子的基序（motif）与大沟的 DNA 相结合而发挥作用。

总之，双螺旋的碱基互补配对结构的 DNA 分子作为生命物质具有重要的生物学意义。它是遗传信息的载体，蕴藏并传递着支配生命活动的指令，是构建生命的草图，也是用于研究生命和改造生命特性的元件。

二、人类基因组 DNA 组成

人类所有 DNA 构成了人类基因组（genome），包括核基因组（nuclear genome）和线粒体基因组（mitochrondrial genome），如图 2-2 所示。核基因组是指每个体细胞核中的父源或母源整套 DNA，也就是说每个体细胞有两套核基因组；每个核基因组的 DNA 约有 3.2×10^9bp。线粒体基因组是指每个线粒体中的闭环双链 DNA，即线粒体 DNA（mitochondrial DNA，mtDNA）（见第九章线粒体遗传病）。基因组中 DNA 序列不同决定了其不同的功能，可分类如下：

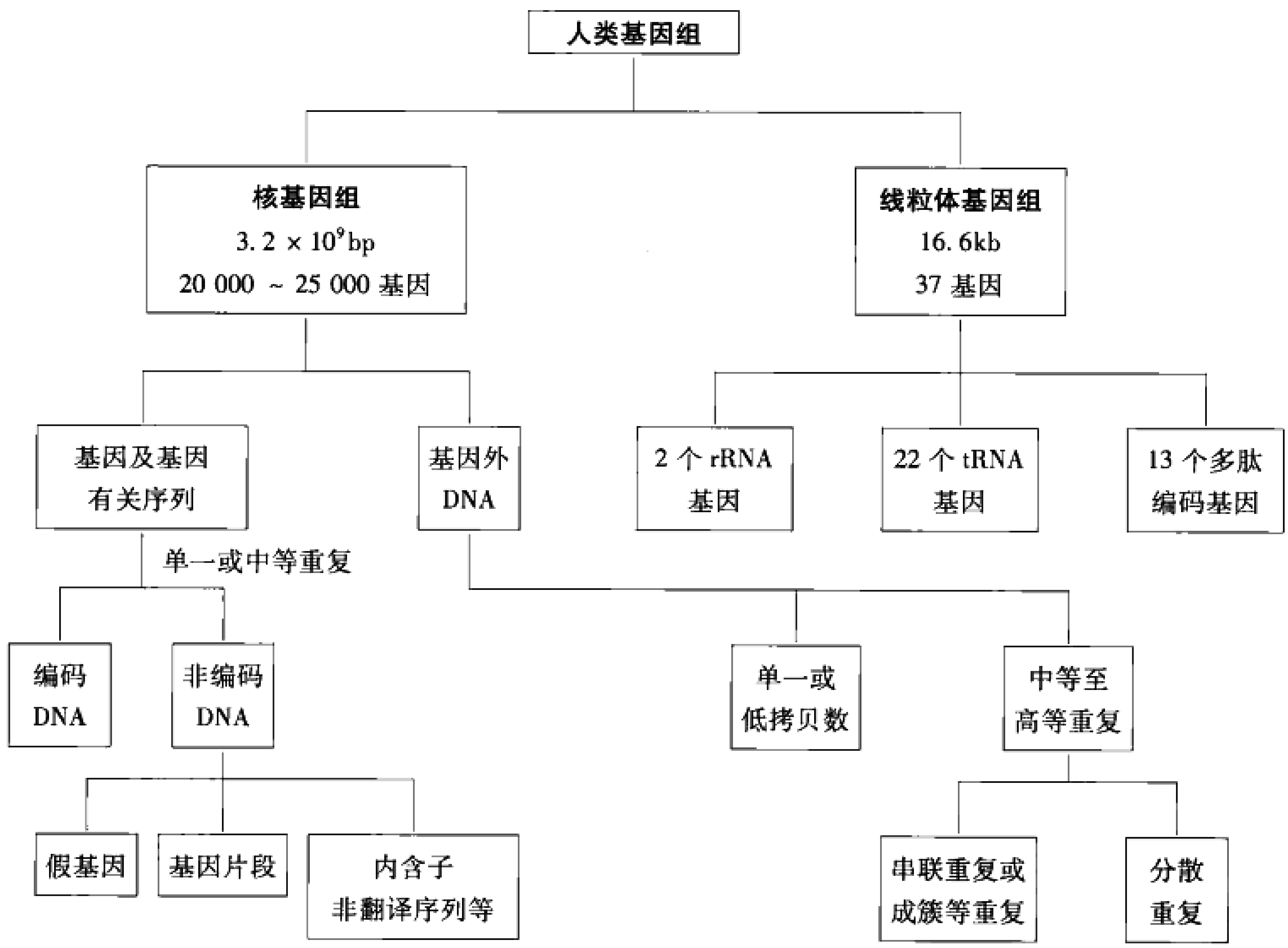


图 2-2　人类基因组的组成

（一）基因序列和非基因序列

基因序列系基因组中决定蛋白质的 DNA 序列，它的一端有起始密码子 ATG，另一端有终止密码子。在起始密码子和终止密码子之间 DNA 序列的长短因不同基因而异，这种 DNA 序列称为开放阅读框（open reading frame，ORF）。也就是说，一个 ORF 相当于一个基因。非基因序列系指基因组中除基因以外的全部 DNA 序列，包括每个基因间的 DNA 序列（intergenic DNA）。

（二）编码序列和非编码序列

编码序列系指编码蛋白质的 DNA 序列，也就是基因中的外显子编码序列；而基因中的内含子是

不编码蛋白质的，因此非编码序列就是基因中的内含子序列以及基因间的序列。

（三）单一序列和重复序列

单一序列（unique sequence）系指在基因组中只出现一次的 DNA 序列即单拷贝 DNA 序列。多数基因为单一序列，非基因序列中也有单一序列；而在基因组中有些基因为多个拷贝数。重复序列（repetitive sequence）系指在基因组中重复出现的 DNA 序列。在人类基因组中约 60% ~70% 是单拷贝或低拷贝的 DNA 序列，包括编码蛋白的基因。30% ~40% 是中度或高度重复 DNA 序列，将基因分隔开。重复序列依其在基因组中重复频率的特征又可分为串联重复序列和分散重复序列。

1. 串联重复序列（tandem repetitive sequence） 指不同长度核苷酸序列的重复单位串联在一起的高度重复序列。一般长度为 2 ~200bp，根据重复单位大小可分为三个主要亚类：卫星 DNA，小卫星 DNA 和微卫星 DNA。

卫星 DNA（satellite DNA）由很大的串联重复 DNA 排列组成，分布在 100kb 至数个 Mb 范围内。重复单位可以是一个简单的短核苷酸序列或一个中等复杂核苷酸序列。卫星 DNA 聚集在染色体着丝粒异染色质区，一般不转录。当 DNA 经氯化铯密度梯度离心时，卫星 DNA 可以与总基因组 DNA 分开，可以看见 DNA 主带之外还有小的卫星带，这是由于卫星 DNA 中 GC 含量少于主带所致。卫星 DNA 的确切功能尚不清楚，但是已知 α 卫星（又称 αDNA）存在于所有染色体上，它是由一个 171bp 重复单位的串联组成并构成着丝粒异染色质的主体，其重复单位通常含有一特异着丝粒蛋白的结合位点。

小卫星 DNA（minisatellite DNA）由重复单位在 6 ~64 个核苷酸的串联重复序列组成，这些序列常在 0.1 ~20kb 范围内，分布于所有染色体的端粒，绝大多数不转录。例如，高可变小卫星 DNA（Hypervariable minisatellite DNA）的重复单位大小不同，但共享一个常见核心序列，GGGCAGGANG（N = 任一核苷酸），其作用仍不清楚，曾报道它是人类细胞中同源重组的热点。小卫星 DNA 序列的另一主要家族是染色体的端粒 DNA，主要由 3 ~20kb 的串联六核苷酸重复单位组成，即 TTAGGG，这是由特异端粒酶加上的，这种简单的重复直接担负着端粒的功能。

微卫星 DNA（microsatellite DNA）由 2 ~6 个核苷酸为重复单位的串联重复序列组成，它们数量多，分散于基因组中，又称为短串联重复序列（short tandem repeat，STR），一般构成染色体着丝粒、端粒和 Y 染色体长臂的染色质区，大多由复制滑动而产生的。双核苷酸重复排列是最常见的类型，约占基因组的 0.5%，出现频率依次是：CA/TG 重复为 1/36kb，AT/TA 重复为 1/50kb，AG/CT 重复为 1/125kb，CG/GC 重复为 1/10Mb，因为 CpG 双核苷酸易于甲基化并继之去氨基。微卫星 DNA 的重要意义尚不清楚，其多态性可作遗传学研究的遗传标记。有些微卫星 DNA 位于基因的编码序列，因为易于复制滑动而常为突变热点。如$(CGG)_n$等三核苷酸重复的动态突变是某些神经肌肉系统疾病的原因。

2. 分散重复序列（interspersed repeated sequence） 这是分布于基因组内散在的重复序列。依其重复序列长短又可分为短分散核元件（short interspersed nuclear element，SINE）和长分散核元件（long interspersed nuclear element，LINE）。

SINE 长度在 100 ~400bp，其拷贝数可达 10^6 以上。在人基因组中 SINE 间平均距离 2.2kb，分散于基因内、基因间或基因簇内，甚至内含子中也有 SINE，但未见于编码区外显子内。*Alu* 序列是其典型代表，也是人类基因组中含量最丰富的重复序列。由 282bp 序列构成，有一限制性内切酶 *Alu I* 的识别序列 AGCT，故得此名，约有 50 万 ~70 万拷贝。*Alu* 序列存在于人类和一些灵长类基因组中，因而可作为人和这些动物基因组的重要标记。

LINE 长度为 5000 ~7000bp，重复拷贝数 10^2 ~10^4 次。如 *Kpn I* 家族可由限制性内切酶 *Kpn I* 切割，分散在基因组中。这些序列构成可转座元件（transposable elements），使 DNA 可在基因组内由一个染色体转移到另一染色体，其功能意义正令人关注。

第二节　基因及其表达与调控

一、基因的基本组成

基因是具有功能的 DNA 序列片段。大多数真核细胞的基因不是连续编码序列，而是由编码序列和非编码序列两部分构成，非编码序列将编码序列隔开，因此把这种基因又称割裂基因（split gene）。人类基因主要由外显子、内含子和侧翼序列组成（图 2-3）。

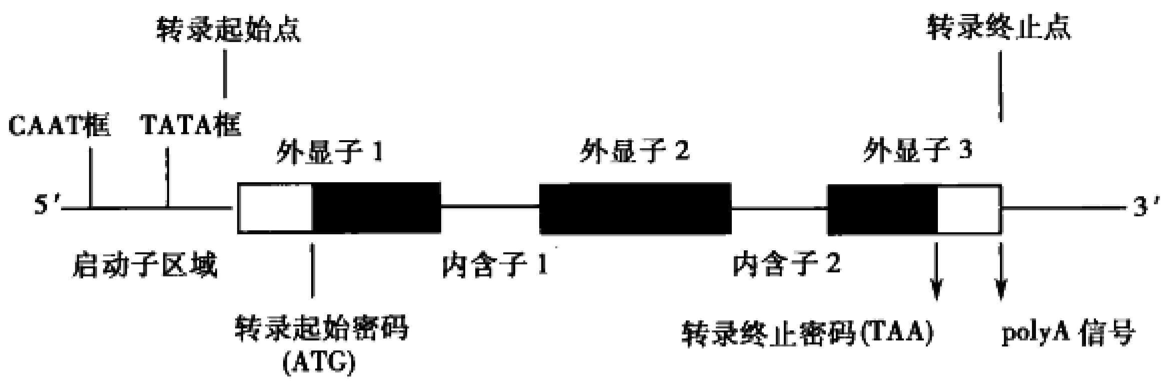


图 2-3　基因结构图解

（一）外显子和内含子

外显子（exon）多数是基因内的编码序列，而内含子（intron）是基因内的非编码序列，内含子在转录成 mRNA 之前被剪切掉，因此不在 mRNA 序列中。基因一般由若干外显子和内含子组成，编码序列的外显子平均长度少于 200bp，非编码序列的内含子平均长度为 3000bp 左右，所以，外显子和内含子组成的基因大小是多样的；没有内含子的基因一般较小，较大的基因一般有较大的内含子，由于内含子的长序列在转录时会耗费时间和能量，因而，对于高表达的基因来说，自然选择利于短的内含子。在每个外显子与内含子的接头部位，都有一高度保守的共有序列，为剪接识别信号，即每个内含子 5′端的两个核苷酸都是 GT，3′端的两个核苷酸都是 AG，这种连接方式称为 GT-AG 法则，是真核细胞中基因表达时剪切内含子和拼接外显子的共同机制。不同基因的外显子和内含子的数目和大小各不相同，一般是基因越大，外显子越多。内含子中还可能存在若干小基因，即基因内基因。

人类基因组中有一些功能相似的基因成簇地排列在一起，称为基因家族（gene family），这些基因可能同时发挥作用，或在不同发育阶段表达，合成某些蛋白。例如人 α 和 β 珠蛋白基因簇，α 珠蛋白基因和附近的 ζ 基因排列在 16 号染色体，共同组成 α 珠蛋白基因簇；β 珠蛋白基因和附近的 ε、$^{G}\gamma$、$^{A}\gamma$ 和 δ 四个基因排列在 11 号染色体上，共同组成 β 珠蛋白基因簇；它们是在胚胎发育不同阶段表达。有一些基因编码相似功能的蛋白，成簇地分布在几条不同的染色体上，又称为基因超家族（gene superfamily），如 *HOX* 基因是由 38 个相关基因组成的四个基因簇，分布在 2、7、12 和 17 号染色体上。此外，在基因家族中还存在一些假基因（pseudogene），是与某些有功能的基因结构相似而不能表达基因产物的基因；可能是在进化中由于基因在复制时，编码序列或调控元件发生突变、或插入了 mRNA 反转录的 cDNA，一般是缺少基因表达所需的启动子序列，故而不能进行转录，例如 α 珠蛋白基因簇中的假基因 ψα 与 α 基因相比，只是没有内含子，可能是 cDNA 插入所致。

（二）侧翼序列

每个基因的 5′和 3′端两侧都有一段不转录的 DNA 序列，称为侧翼序列（flanking sequence）。启动子在 5′端，终止子和多聚腺苷酸信号在 3′端；增强子可能在 5′或 3′端。这些 DNA 序列是与基因转录有关的重要作用元件，对基因的转录表达起重要调控作用。

1. 启动子　启动子（promoter）由一组短序列元件簇集在一个基因编码序列的上游构成一个启动

子，一般位于基因转录起始点上游100～200bp范围；转录因子与启动子结合后，激活RNA聚合酶，在特定位置起始RNA合成。启动子主要包括下列一些作用元件：

（1）TATA框（TATA box）：人类许多基因在转录起始点5′端上游－25～－30bp处有一段高度保守序列，由7个碱基组成，即TATAA（T）AA（T），其中两个碱基可以变化。TATA框能与转录因子TFⅡ结合，再与RNA聚合酶Ⅱ形成复合物，准确识别转录起始点，启动基因转录。

（2）CAAT框（CAAT box）：位于转录起始点5′端上游－70～－80bp有一段保守序列，由9个碱基组成，即GGC（T）CAATCT，其中有一个碱基可以变化。CAAT框能与转录因子CTF结合，提高转录效率。

（3）GC框（GC box）：有一些基因没有TATA框和CAAT框，但是存在富含G和C核苷酸的序列；GC框是由GGCGGG组成，能与转录因子Sp1结合，促进转录过程。

2. 增强子　增强子（enhancer）是一个短序列元件，特异性地与调节蛋白结合，在启动子和增强子之间形成DNA环，使得增强子的结合蛋白质与启动子的结合蛋白质相互作用、或者与RNA聚合酶相互作用，增强基因的转录活性。增强子与启动子有所区别，启动子位于基因上游，起始点相对恒定；增强子可以在基因的任何位置，而且其功能与位置和序列方向无关，可以是5′→3′方向，也可以是3′→5′方向。

3. 沉默子　沉默子（silencer）是与增强子具有相似性质，但是抑制特定基因转录活性的调节元件。

4. 终止子　终止子（terminater）是由AATAAA和一段回文序列组成，AATAAA是多聚腺苷酸（polyA）附加信号，回文序列转录后形成发夹结构，阻碍RNA聚合酶继续移动，转录终止。

二、基因的表达

基因表达（gene expression）是DNA序列的遗传信息通过转录（transcription）产生的mRNA再经翻译（translation）最终生成蛋白质的过程。基因表达遵循共线性原理（colinearity principle），即DNA的线性核苷酸序列以碱基三联体（base triple）形式被转录为RNA的线性核苷酸序列，RNA以密码子（codon）形式被解码形成特定多肽的线性氨基酸序列，这种DNA→RNA→蛋白质的信息传递原则称为中心法则。由于发现了反转录酶，能够以RNA序列为模板合成DNA，因此遗传信息并非单方向传递，可以是DNA⟷RNA→蛋白质。

（一）转录

转录（transcription）是指以DNA双链中的一条链为模板，以ATP、CTP、GTP、UTP作为原料，在RNA聚合酶催化下，按碱基互补方式合成RNA单链的过程。转录发生在细胞核中，转录方向为5′→3′，转录产物RNA的碱基序列与DNA模板链互补，同非模板链一致，只是把T换成了U。因此经常把非模板链称为有义链（sense strand），模板链称为反义链（antisense strand）。

真核细胞中只有一小部分DNA依其需要而被转录，DNA转录单位无规律地分布在基因组DNA中。转录产物有三种，信使RNA（messager RNA，mRNA），核糖体RNA（ribosomal RNA，rRNA）和转运RNA（transfer RNA，tRNA），它们分别是由RNA聚合酶Ⅱ，RNA聚合酶Ⅰ和RNA聚合酶Ⅲ催化合成的；只有mRNA将遗传信息传递给蛋白质，成熟的mRNA是原始RNA转录本经过一系列的加工而成为合成多肽链的模板。加工一般包括剪接、加帽、加尾等过程（图2-4）：

1. 剪接（splice）　基因的外显子和内含子转录成的原始RNA转录本称为异质核RNA（heterogenous nuclear RNA，hnRNA），剪接是将内含子的RNA序列切掉，将外显子的RNA序列拼接起来的过程。剪接发生在外显子和内含子交接处的GT和AG；剪接起始的GT和相邻的保守序列组成了剪接供体位点（splice donor site），剪接终止的AG和相邻的保守序列组成了剪接受体位点（splice receptor site），在接近内含子末端有一个保守序列称分支部位（branch site），一般位于AG上游40核苷酸处，这些序列构成了剪接信号（图2-5），被细胞核内的小核RNA蛋白（snRNP）识别并结合，形成剪接体

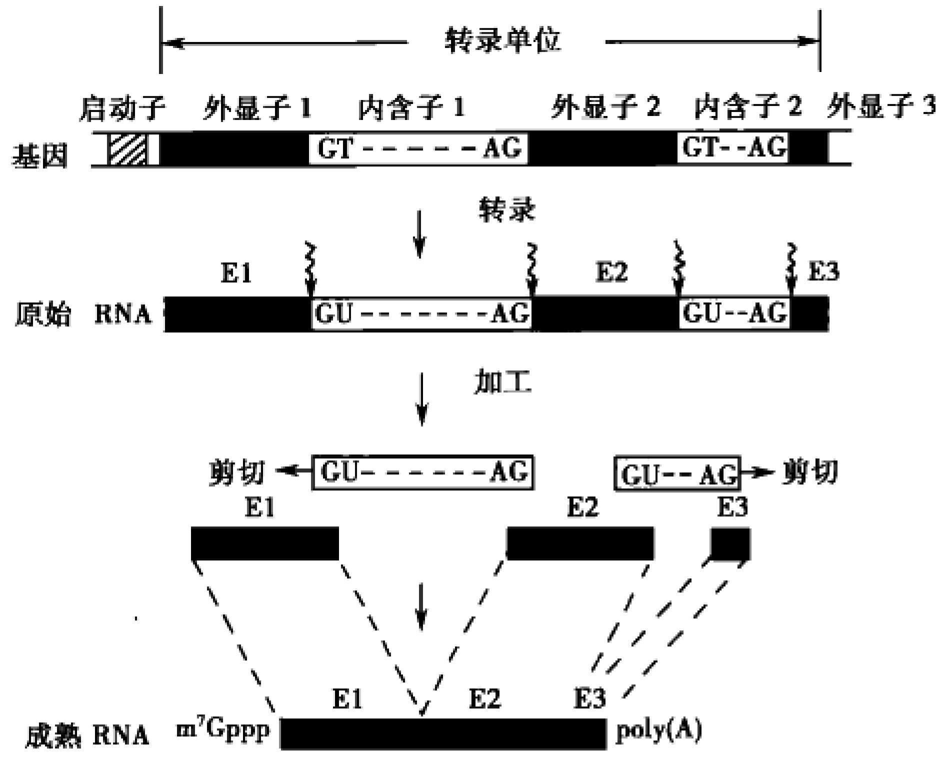


图 2-4 RNA 加工图解

(splicesome)并切除内含子。snRNP 是由小核 RNA(共 5 种,即 $snRNAU_1$、$snRNAU_2$、$snRNAU_4$、$snRNAU_5$ 和 $snRNAU_6$)和特定蛋白质组成,snRNP 和 RNA 转录本的 RNA-RNA 碱基配对,决定了剪接反应的特异性。

2. 加帽(capping) 是指在 RNA 转录本 5′端连接上一个 7-甲基鸟苷酸,封闭了 RNA 的 5′端,称为加帽。加帽的作用在于保护 RNA 转录本避免被磷酸酶和核酸酶消化而增强 mRNA 的稳定性,有利于 mRNA 从细胞核运到细胞质及被细胞质中的核糖体所识别。

3. 加尾(tailing) 是指在加帽的同时,RNA 转录本 3′端在腺苷酸聚合酶作用下,经多聚腺苷酸化(polyadenylation)附加大约 200 个腺苷酸的长链,即多聚腺苷酸(polyA)尾。加尾是在 3′端非编码区一个 6 核苷酸信号 AAUAAA 的下游 15 ~ 30bp 的部位加上 polyA。polyA 促进 mRNA 从细胞核向细胞质的转运,稳定细胞质中 mRNA 分子以及有利于核糖体识别 mRNA 而促进翻译。

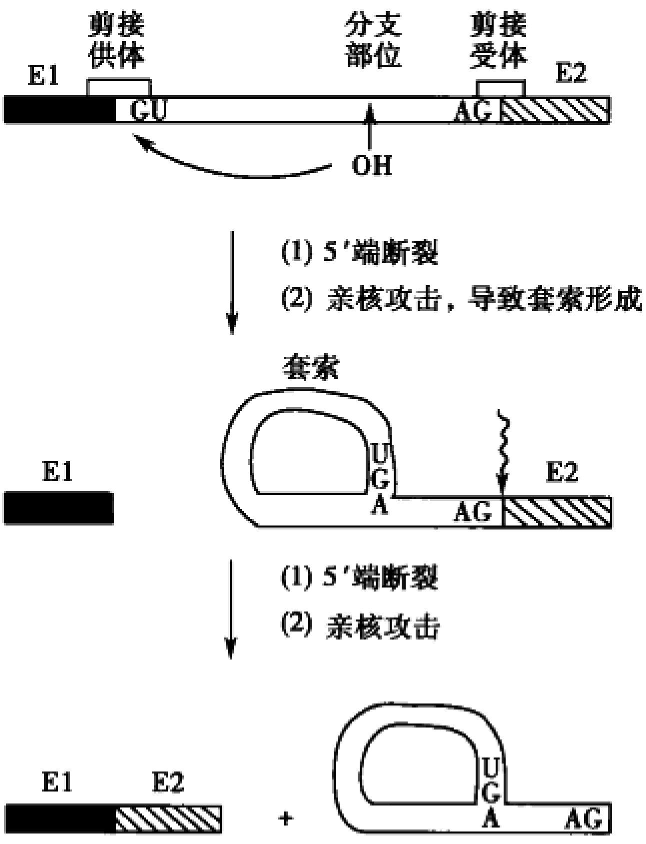


图 2-5 RNA 剪接机制

(二) 翻译

翻译(translation)是指 mRNA 转译成氨基酸序列的过程。成熟的 mRNA 从细胞核进入细胞质,由核糖体阅读 mRNA 所携带的信息,指导特异的多肽合成。通常 mRNA 的中间序列被翻译成氨基酸,而 5′端和 3′端是非翻译区(5′UTR 和 3′UTR),多数为第一和最后外显子序列,其中包括了 5′端的加帽和 3′的加尾序列。

1. 翻译过程 多肽链的合成是在 mRNA、tRNA 和核糖体协同作用下进行的。核糖体是一个 rRNA-蛋白质复合物,由 60S 和 40S 两个亚基组成。核糖体的小亚基识别 mRNA 5′端的“帽子”,沿着 mRNA 序列移动到第一个起始密码子 AUG;然而 AUG 只有嵌入在起始密码子识别序列 GCCPuCCAUGG 时,才能够作为一个起始密码子被有效识别;此序列中最重要的是 AUG 密码子之后的 G,以及在它之前第三个核苷酸的嘌呤(Pu),最好是 A。在识别起始密码子后,多种 tRNA 携带特定的氨

基酸，tRNA 上的反密码子逐一识别 mRNA 上互补的密码子，核糖体的大亚基结合小亚基开始精确地合成肽链。整个过程按进位、转肽、移位和脱落等步骤不断重复进行直到终止密码子（UAA，UAG，UGA），使多肽链从核糖体上释放出来。翻译过程并非是单个核糖体在一个 mRNA 分子上进行，而是有好几个甚至几十个核糖体在同一条 mRNA 分子上进行翻译，形成多聚核糖体，可按不同进度翻译成多条的多肽链。mRNA 的 5′端对应于蛋白的氨基末端 NH_2，而其 3′端则对应于蛋白质的羟基末端（COOH）。

2. 遗传密码的兼并性　DNA 分子的每三个碱基组成一个遗传密码，对应于 RNA 分子的密码子。核酸分子中有 4 种碱基，可以组成 64（4^3）个密码子。而氨基酸只有 20 种，每个氨基酸平均有三个密码子编码；蛋氨酸只有一个密码子，亮氨酸和丝氨酸分别有 6 个密码子，这种不同密码子编码同一氨基酸的特性称为遗传密码子的兼并性（degeneracy）（表 2-1）。mRNA 的密码子有 64 个，而细胞质 tRNA 的反密码子有 30 个，线粒体 tRNA 的反密码子有 22 个，但是在翻译过程中 tRNA 仍然有效地运送氨基酸。关于密码子和反密码子的互补配对有一个摇摆假说（wobble hypothesis），即第一和第二碱基遵循 A-U 和 G-C 规律，第三碱基可以发生“摇摆”出现 G-U 配对。

表 2-1　遗传密码

第一个核苷酸（5′端）	第二个核苷酸				第三个核苷酸（3′端）
	U	C	A	G	
U	UUU 苯丙 Phe，F	UCU	UAU 酪 Tyr，Y	UGU 半胱 Cys，C	U
	UUC	UCC 丝 Ser，S	UAC	UGC	C
	UUA 亮 Leu，L	UCA	UAA 终止	UGA 终止	A
	UUG	UCG	UAG	UGG 色 Trp，W	G
C	CUU	CCU	CAU 组 His，H	CGU	U
	CUC 亮 Leu，L	CCC 脯 Pro，P	CAC	CGC 精 Arg，R	C
	CUA	CCA	CAA 谷酰 Gln，Q	CGA	A
	CUG	CCG	CAG	CGG	G
A	AUU	ACU	AAU 天酰 Asn，N	AGU 丝 Ser，S	U
	AUC 异亮 Ile，I	ACC 苏 Thr，T	AAC	AGC	C
	AUA	ACA	AAA 赖 Lys，K	AGA 精 Arg，R	A
	AUG 蛋 Met，起始	ACG	AAG	AGG	G
G	GUU	GCU	GAU 天冬 Asp，D	GGU	U
	GUC 缬 Val，V	GCC 苏 Thr，T	GAC	GGC 甘，Gly，G	C
	GUA	GCA	GAA 谷 Glu，E	GGA	A
	GUG	GCG	GAG	GGG	G

3. 翻译后修饰　初始翻译的多肽链需要进一步加工修饰，才能形成具有一定空间结构和活性的蛋白质。翻译后的修饰主要有脱甲酰基、乙酰化、磷酸化、糖基化和链切割等，以及两条以上肽链间的连接和进一步折叠形成特定的空间构象等。例如输送到溶酶体、高尔基体、浆膜或细胞分泌的蛋白质都需要糖基化，即在某氨基酸侧链附着寡糖成为糖蛋白；膜蛋白的脂酰基团起膜固定作用；像血浆蛋白、多肽激素、神经多肽、生长因子等多肽链需剪切后才能成为有活性的产物。所有分泌型多肽都是先合成蛋白质前体，其 N-末端的信号序列引导蛋白质前体定位于膜上，然后信号肽被切除。还有一些蛋白质含有其他的信号序列，例如转录因子、DNA 聚合酶、RNA 聚合酶的细胞核定位信号，使蛋白质从细胞质进入细胞核，发挥生物学活性。

三、基因表达的调控

人体所有细胞都含有完整的基因组，但在特定组织中只有部分基因表达，而不同基因是在不同的时空进行表达的。细胞类型的区别并非所含基因组的不同，而是基因表达差异所致。基因若在不当的时空表达或其产物量的异常都能引起疾病。因此，认识基因表达调控对了解人类生命本质以及疾病发生机制都是十分重要的。在多数细胞中都表达的基因称管家基因（housekeeping gene），例如核糖体、染色体、细胞骨架的相关蛋白基因。一般来说，基因表达受如下因素影响：①基因转录成RNA的速率；②RNA的加工；③mRNA的稳定和降解速率；④mRNA翻译成蛋白质的速率；⑤蛋白质翻译后修饰；⑥蛋白质的稳定和降解速率。基因表达可能涉及所有这些调控机制，然而初始转录水平的调节是最为关键的。

（一）组蛋白乙酰化和DNA甲基化

细胞核中的染色质可以分为无转录活性和有转录活性两种，无转录活性的染色质通常呈高密度，在细胞周期的S期晚复制，与组蛋白紧密结合，DNA甲基化程度高。有转录活性的染色质较松散，在S期早复制，与组蛋白结合较弱，组蛋白乙酰化，DNA甲基化相对少，特别是基因启动子区域。组蛋白乙酰化的程度是影响基因转录的一个重要因素，组蛋白乙酰转移酶将乙酰基加在组蛋白N端的赖氨酸上，使N端形成一个尾巴，突出在核小体外；由于乙酰化的组蛋白对DNA的亲和力减低，造成染色质开放结构，更适合基因表达。组蛋白的去乙酰化与基因启动子中CpG岛的甲基化相关，抑制基因转录。当甲基化的CpG序列与甲基化CpG结合蛋白（MeCP2）结合后，可以被转录抑制因子和组蛋白去乙酰酶组成的复合体识别，使染色质呈紧密结构。

（二）顺式作用元件和转录因子

基因启动子中有一些保守序列能与转录因子特异性结合，调节基因转录，这些元件称为顺式作用元件（cis-acting element）。其功能的发挥限于基因5′端侧翼序列的所在位置，调节邻近基因的表达。真核细胞中除了启动子中的TATA框，CAAT框外，还有其他一些顺式作用元件，通过DNA-蛋白质相互作用，特异性地调控基因表达。例如类固醇激素对基因的表达调节是由该基因启动子中的激素反应元件与激素-激素受体复合体的特异性结合决定的。

真核细胞中的RNA聚合酶本身不能启动转录，必须有许多转录因子特异结合在基因上游的顺式作用元件后才激活RNA聚合酶，从转录起始点开始合成RNA。通常把转录因子称为反式作用因子（trans-acting factor），“trans”意味着转录因子转移到它们作用的位置，对应于“cis”。在已知众多的转录因子的结构中都有一些相似的结构域基序，这些基序就是蛋白质与DNA顺式作用元件特定序列相结合的部位。转录因子不仅与DNA靶序列相互作用，而且它们之间也相互作用，正是这些作用决定了高等生物和人类发育过程中所必须而复杂的组织特异性基因表达。转录因子根据与DNA结合的结构域基序分为四种，最常见的是螺旋-转角-螺旋（helix-turn-helix）蛋白质，由一个氨基酸短链连接两个α螺旋结构构成，其他三种为锌指蛋白（zinc finger），亮氨酸拉链蛋白（leucine zipper）和螺旋-环-螺旋蛋白（helix-loop-helix）。

（三）剪接与多聚腺苷酸化

一个基因转录本的剪接改变可以形成许多异构蛋白，例如组织特异性的异构蛋白，膜结合的和可溶性的异构蛋白等。参与剪接调控的主要调节因子是RNA结合蛋白的SR家族（C端含丝氨酸和精氨酸）和一些snRNP蛋白，这些蛋白能促进剪接体装配的每一环节结合于剪接增强子序列。在许多基因中，3′UTR可能有两个或更多的多聚腺苷酸信号，不同多聚腺苷酸化的转录本表现了组织的特异性，例如降钙素基因（*CALC*）在甲状腺组织表达为降钙素，是一个与循环钙离子稳定有关的激素；但在下丘脑组织中却表达为降钙素相关肽（CGRP），具有神经调节和营养活性。

第三节　遗传信息的传递

一、DNA 复制

DNA 复制就是 DNA 合成过程，DNA 通过复制把储存的遗传信息随着细胞的分裂传递给子细胞。DNA 复制的主要特点是：

（一）半保留复制

半保留复制是指 DNA 复制过程中，DNA 双链被解旋酶分成两条单链，每条单链都能指导合成一条互补链，形成两个子 DNA 双链。由于每个子 DNA 的两条核苷酸链一条来自亲代 DNA，另一条是新合成的 DNA，因此复制过程是半保留的（图 2-6）。

（二）半不连续复制

在复制过程中，DNA 以脱氧三磷酸核苷（dATP、dCTP、dGTP、dTTP 统称 dNTP）为原料，在 DNA 聚合酶催化下合成新链。由于 dNTP 原料只能连接在多核苷酸链游离的 3′碳原子的羟基 OH 上，所以 DNA 复制方向是 5′→3′进行。DNA 复制是从特异复制起点开始，进行双向复制。以 3′→5′DNA 链为模板，按 5′→3′方向连续复制，速度较快，复制较早完成，合成的链称为前导链。以 5′→3′DNA 为模板，按 5′→3′方向进行复制，先合成一段 DNA 片段，约 100～1000 碱基，称为冈崎片段，DNA 连接酶将这些冈崎片段连接起来，形成完整单链，这个复制过程完成较晚，合成的链称为后随链。DNA 复制是一条链为连续复制，另一条链为不连续复制，因此复制是半不连续的。

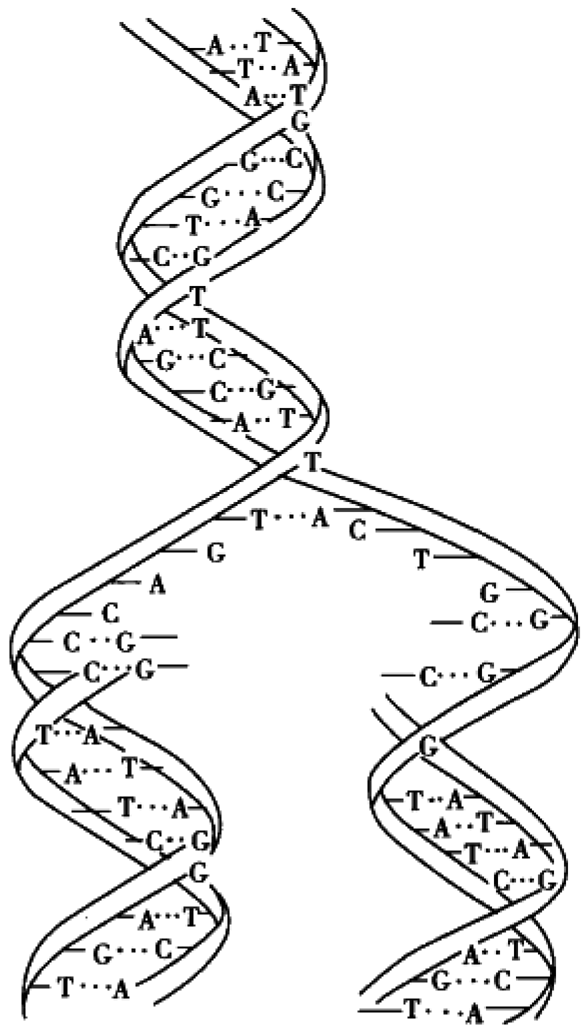


图 2-6　DNA 半保留复制

（三）复制子

真核细胞的 DNA 复制有许多复制起点，在一个复制起点上进行的 DNA 复制区段为一个复制单元称为复制子（replicon）。相邻的复制起点相距约 30～50kb，在复制过程中，许多复制子同时进行复制，从起点开始双向进行，在两侧形成复制叉（replication fork），相邻复制叉移动至汇合处相连，复制终止（图 2-7）。

二、细胞分裂

生长和繁殖是生物体的普遍生命现象。生长是受精卵不断通过有丝分裂增加细胞数量，同时进行分化形成不同类型的体细胞并增大体积。繁殖是生殖细胞通过减数分裂产生成熟的精子和卵子，再经受精形成受精卵，而后生长发育产生可延续的后代。细胞分裂分两类：有丝分裂和减数分裂。

（一）有丝分裂

有丝分裂（mitosis）是体细胞分裂的基本方式，是一个细胞分裂产生两个在遗传上与亲代完全相同的细胞的过程。有丝分裂使一个受精卵发育成为一个约有 10^{14} 个细胞的人体，所有这些细胞与原始单细胞所含有的遗传信息完全相同。一个人在一生中要经过 10^{17} 次有丝分裂，每个子代细胞接受全部的 46 条染色体。细胞从上一次有丝分裂结束到下一次有丝分裂完成的全过程称为细胞周期（cell cycle），细胞周期包括了细胞间期和分裂期（图 2-8）。

1. 细胞间期　间期（interphase）是两次细胞分裂之间的相隔时期。这是 DNA 合成和细胞代谢活

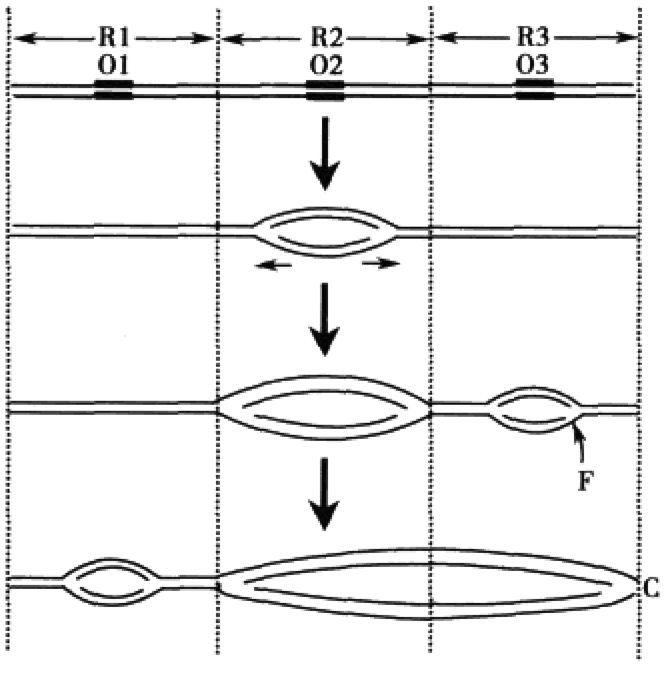


图 2-7 DNA 的双向复制

动旺盛的时期，对于快速分裂的细胞约为 16～24 小时。依据 DNA 合成和染色体形态的变化，间期又分为：

G_1 期（G：gap）：主要是 RNA 和蛋白质的合成，DNA 尚未复制，但为 DNA 合成进行物质和能量的准备，如合成 DNA 前身物和一些酶类。该期在不同类型细胞的时间各异，停止分裂的细胞如神经元常常在 G_1 期静止，进入 G_0 期。

S 期（S：synthesis）：主要进行 DNA 合成，DNA 总量增加一倍，染色体中的两条染色单体已经形成。同时还进行组蛋白和非组蛋白等染色体蛋白的合成。

G_2 期：有少量 DNA 和蛋白质合成，为细胞进入分裂期准备物质条件，如微管蛋白等有丝分裂期成分即染色体凝集因子的合成。

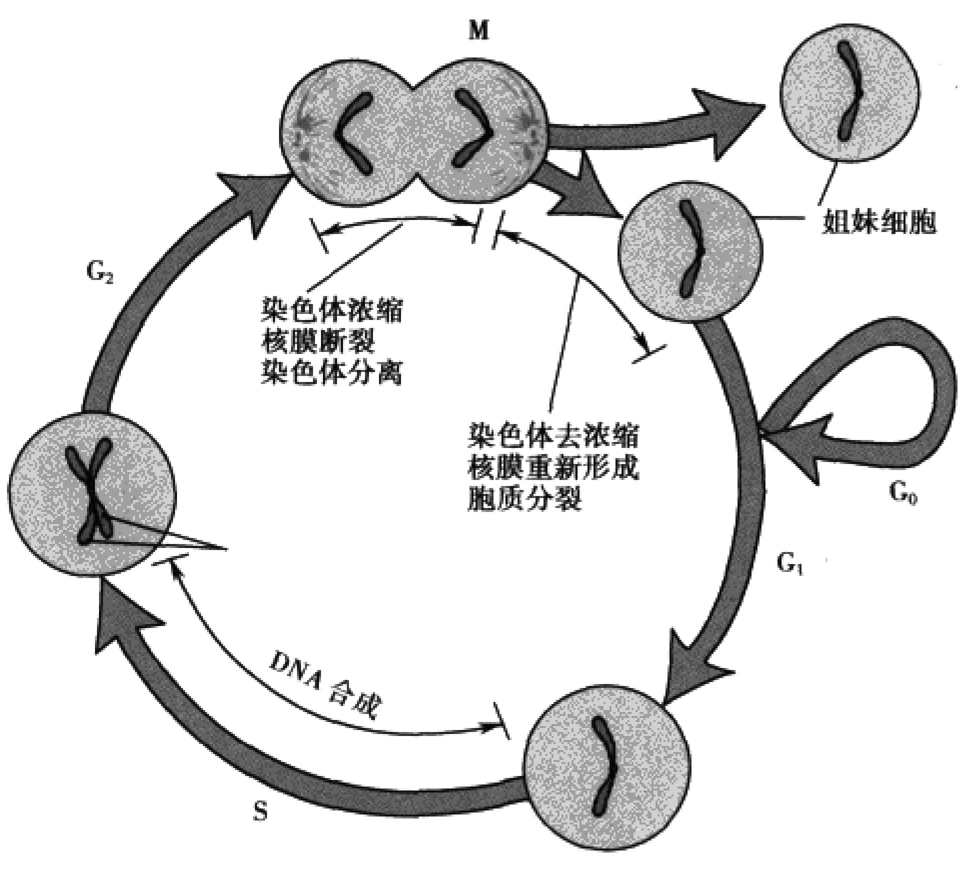


图 2-8 细胞周期

此时已有两套完整的二倍体染色体，分裂期是染色体分离和核分裂的过程。

2. 分裂期　间期细胞经过复制准备的两套染色体迅速分离并传送到两个子细胞即有丝分裂期也称 M 期（M：mitosis）。分裂期时间短，完成快。依据分裂过程中染色体的变化，分为连续的四期：前期、中期、后期和末期（图 2-9）。

前期（prophase）：细胞核内染色质变成纤细而屈曲的细丝，通过螺旋化和折叠逐渐缩短和变粗，形成具有明显形态结构和一定数目的染色体，核仁核膜同时消失。

中期（metaphase）：染色体最终凝缩为典型的形态结构，每条染色体具有两条染色单体（chromatid），它们仅在着丝粒（centromere）处相连，此时中心粒分离到细胞的两极，纺锤体形成并与每条染色体的着丝粒相连，染色体在细胞中央排列成赤道板。

后期（anaphase）：每条染色体的着丝粒纵裂，两条染色单体分别由纺锤丝牵引而移向细胞两极。移向细胞两极的两组染色体形态和数目完全相同。细胞膜从中间开始凹陷。

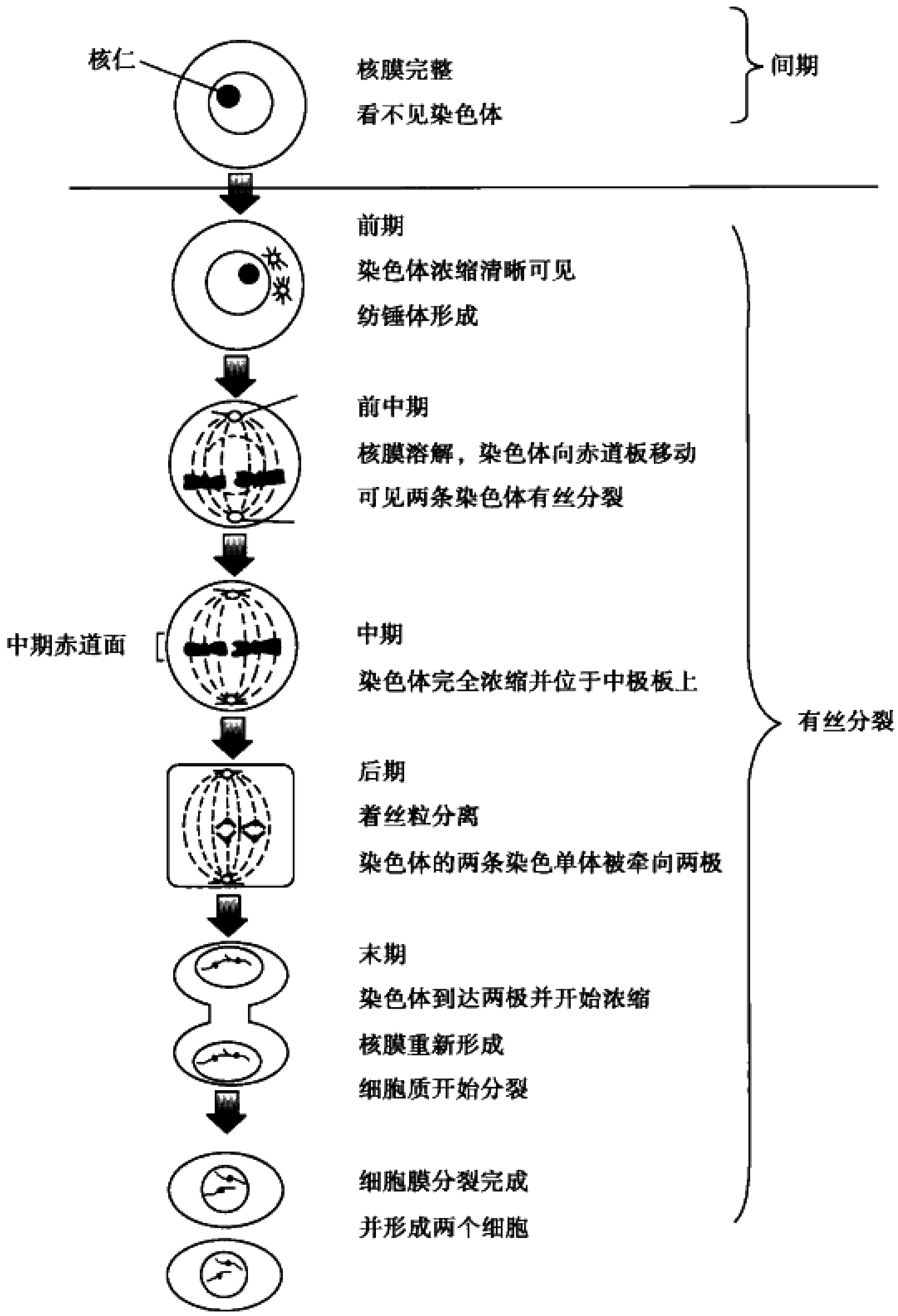


图 2-9 有丝分裂图解

末期(telophase)：两组染色体集中于细胞两极并又逐渐解螺旋形成染色质，同时核膜和核仁重新出现。细胞膜凹陷最终将原来的一个细胞分割成两个子细胞。

经过上述一次有丝分裂过程，细胞分裂了一次，DNA 复制了一次，染色体同时也分裂了一次，并且被平均分配到两个子细胞中。因此每个子细胞都含有与原来亲代细胞相同的染色体数，这样就保证了有丝分裂产生的子细胞都保持了与亲代细胞相同的全套遗传信息，从而保证了机体所有细胞的染色体数目的恒定性。

（二）减数分裂

减数分裂(meiosis)是配子发生过程中产生精子和卵子的一种特殊细胞分裂，在生殖细胞成熟过程中 DNA 复制一次后，细胞连续分裂两次，结果形成的配子中染色体数目减少了一半，只含有单倍体的染色体(n)。由于这两次连续分裂发生在配子形成的成熟期，故减数分裂又称为成熟分裂(maturation division)。减数分裂的过程包括了减数分裂Ⅰ和减数分裂Ⅱ(图 2-10)。

1. 减数分裂Ⅰ　细胞在间期Ⅰ完成了 DNA 的复制后，将进入第一次减数分裂。

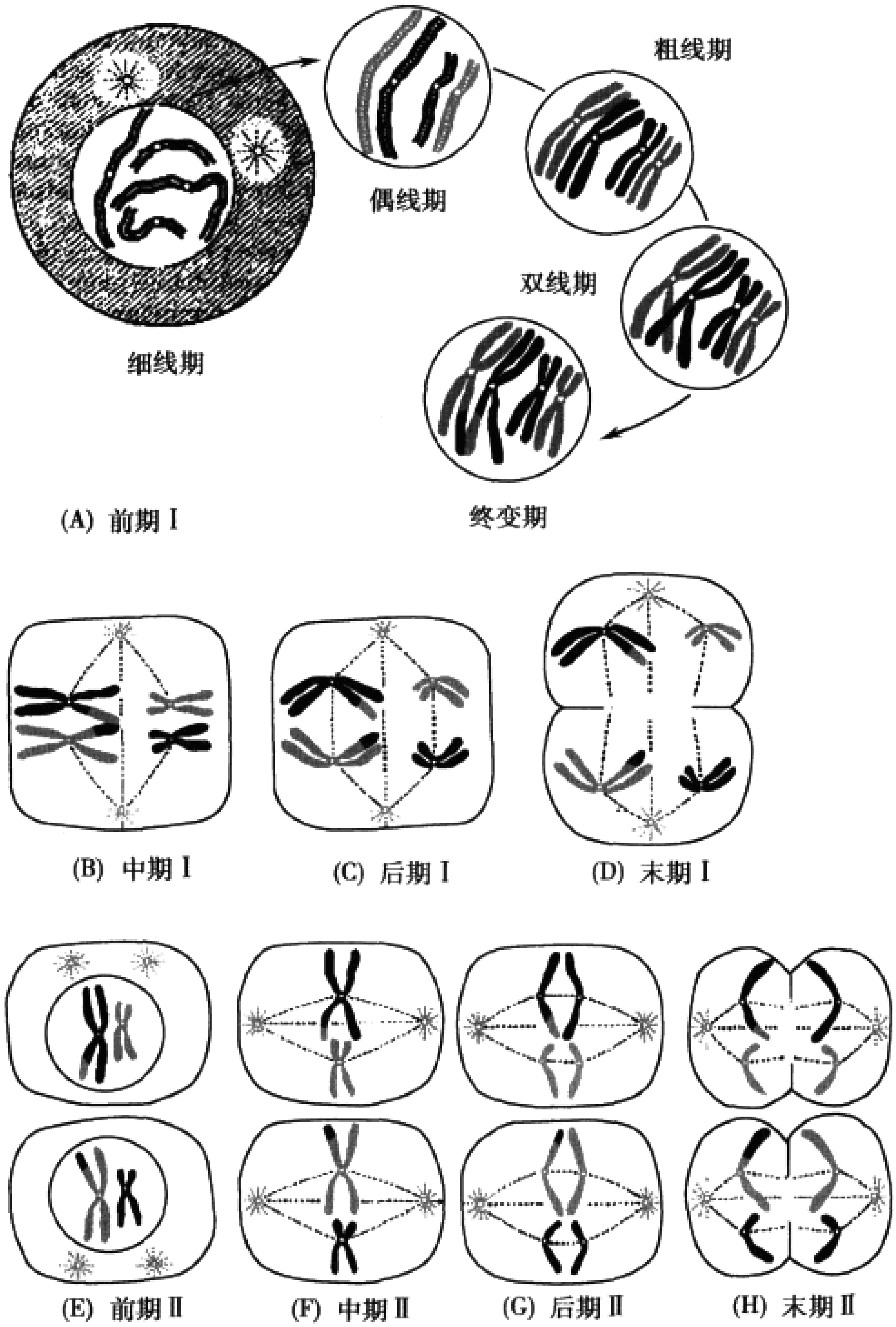


图2-10 减数分裂图解

(1) 前期Ⅰ:减数分裂Ⅰ的前期是一个复杂的变化过程,比有丝分裂的前期要长得多,而且染色体的形态变化比较复杂;可以分为以下几个阶段:

1) 细线期(leptotene stage):染色体呈细丝状,相互交织呈网状,染色质开始凝缩。

2) 偶线期(zygotene stage):同源染色体发生配对,这一精密的排列结合称为联会(synapsis),它构成了遗传物质交换的基础。联会的结果是每对染色体形成一个紧密相伴的二价体(bivalent)。

3) 粗线期(pachytene stage):染色体进一步螺旋化,变粗变短,在光学显微镜下看到的每一条染色体包含两条染色单体,互称为姐妹染色单体。每个二价体都是由同源染色体组成。这样一个二价体有四条染色单体,称为四分体(tetrad)。同源染色体的染色单体之间互称为非姐妹染色单体。在此期间,非姐妹染色单体横向断裂,并在断裂处发生交换(crossing-over)。

4) 双线期(dipleotene stage):同源染色体之间的联会解体,同源染色体相互排斥而趋向分离,使互换后的染色体之间出现交叉(chiasma)。一般认为,交叉是非姐妹染色单体交换的表现形式。

5) 终变期(diakinesis stage):染色体进一步变粗变短。核仁、核膜消失,纺锤体开始形成。

(2) 中期Ⅰ:二价体排列在赤道板上,纺锤丝与着丝粒相连,使两条同源染色体随机分布在赤道板的两侧,这种排列组合意味着人类23对染色体在单倍体细胞中可有2^{23}(8 388 608)种组合方式,这

种现象符合孟德尔分离率和自由组合率，而保证了个体间遗传的多样性。

（3）后期Ⅰ：二价体中的两条同源染色体被纺锤丝牵引彼此分开，分别向两极移动。

（4）末期Ⅰ：染色体到达两极，解旋伸展。核仁，核膜重新出现。

细胞质分开形成两个子细胞。每个子细胞都具有 n 个二分体（dyad），即 n 个经过复制的染色体。

2. 减数分裂Ⅱ　是减数分裂Ⅰ结束后所发生的与有丝分裂相似的细胞分裂。当生殖细胞完成减数分裂Ⅰ，两个子细胞进入间期Ⅱ，但间期很短，而且不再进行 DNA 复制，往往有些生物没有间期Ⅱ，直接进入减数分裂Ⅱ的前期Ⅱ。

（1）前期Ⅱ：减数分裂Ⅱ的前期和细胞的有丝分裂基本上相同，核仁、核膜消失，每个细胞只有 n 个染色体。每个染色体由两条染色单体所组成。

（2）中期Ⅱ：各个染色体排列在赤道板上。

（3）后期Ⅱ：每个染色体的着丝粒纵裂，姐妹染色单体分开，形成两个单分体（monad），并且分别移向两极。

（4）末期Ⅱ：两极各得到了一组单分体，形成了子细胞核，结果每个细胞各具有 n 个单分体，即 n 条染色体，成为完全的单倍体细胞。

减数分裂后染色体数目减半，产生单倍体（n）的生殖细胞（精子和卵细胞），受精后形成的受精卵中染色体又恢复为二倍数（2n）；经过有丝分裂形成的体细胞中都具有二倍数（2n）的染色体，世世代代细胞的染色体数目不变，维持了各个物种的染色体的恒定。同时，减数分裂的过程中同源染色体进行配对，非同源染色体又可以进行自由组合，同源染色体之间可以进行交换，从而产生了遗传物质的重新组合，形成了生物个体的多样性。

三、遗传的基本规律

孟德尔（Mendel G，1822～1884）是遗传学的奠基人，他于 1865 年通过豌豆杂交实验提出了遗传因子（genetic factor）分离律和自由组合律；1910 年摩尔根（Morgan TH，1866～1945）利用果蝇进行杂交实验发现了连锁和交换律，从而奠定了遗传学的理论基础。

（一）分离律

生物在生殖细胞形成过程中，同源染色体分离，分别进入不同的生殖细胞，即每个生殖细胞只有亲代成对的同源染色体中的一条；位于同源染色体上的等位基因也随之分离，生殖细胞只含有两个等位基因中的一个；对于亲代，其某一遗传性状在子代中有分离现象；这就是分离律。

（二）自由组合律

生物在生殖细胞形成过程中，非同源染色体之间是完全独立的，可分可合，随机组合；这就是自由组合律。

（三）连锁和交换律

同一条染色体上的基因彼此间是连锁在一起的，构成了一个连锁群（linkage group）；同源染色体上的基因连锁群并非固定不变。在生殖细胞形成过程中，同源染色体在配对联会时发生交换，使基因连锁群发生重新组合；这就是连锁和交换律。

同源染色体上的两对等位基因之间的交换与连锁基因间的距离有关，相距愈远，发生交换的机会愈大。基因在染色体上的距离用厘摩（centi Morgan，cM）为单位表示，1% 交换率为 1 个厘摩。

第四节　基因突变与修复

一、突变

人类基因组既要保持相对稳定又要有所变化，基因组的 DNA 若一成不变，也就不会有进化，这种变化主要是 DNA 序列的突变。自然界中会发生 DNA 自发突变，但突变率很低，高等生物的自发突变

率为 $1\times10^{-10}\sim1\times10^{-5}$，即在10万~100亿个配子中可能有一个突变。人类的突变率约为百万分之一，大多数会自发进行DNA复制和修复。通过物理、化学和生物因素损伤DNA称诱发突变。DNA突变可以发生在编码序列或非编码序列，可以发生在体细胞，不传递给子代，也可以发生在配子，传递给子代。基因突变既是遗传变异的主要来源，也是进化过程中选择的对象，突变是进化的动力。有害的基因突变构成了群体的遗传负荷，是导致某种表型异常的直接原因，也可能导致个体对于疾病易感性的增加。常见的突变是单个碱基的替换、缺失或插入。也可出现多个碱基的变化，如大小不同片段的缺失和插入。突变类型如下(图2-11)：

正常	AGT	CAG	CAG	CAG	TTT	TTA	CGT	AAC	CCG	… DNA
	Met	Gln	Gln	Gln	Phe	Leu	Arg	Asn	Pro	氨基酸
同义突变	AGT	CAG	CAG	CAG	TTT	TTG	CGT	AAC	CCG	… DNA
	Met	Gln	Gln	Gln	Phe	Leu	Arg	Asn	Pro	氨基酸
错义突变	AGT	CAG	CAG	CAG	TTT	TCA	CGT	AAC	CCG	… DNA
	Met	Gln	Gln	Gln	Phe	Ser	Arg	Asn	Pro	氨基酸
无义突变	AGT	CAG	CAG	CAG	TTT	TGA	CGT	AAC	CCG	… DNA
	Met	Gln	Gln	Gln	Phe	终止	Arg	Asn	Pro	氨基酸
移码突变	AGT	CAG	CAG	CAG	TTT	TAC	GTA	ACC	CG	… DNA
(一个碱基缺失)	Met	Gln	Gln	Gln	Phe	Tyr	Val	Thr	Arg	氨基酸
动态突变	AGT	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	… DNA
(三核苷酸重复)	Met	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	氨基酸

图2-11 突变类型

(一) 点突变

点突变(point mutation)是指一个碱基被另一个碱基所替代，又称碱基替换(substitution)；这是最常见的突变。嘧啶之间或嘌呤之间的替换称为转换(transition)；嘌呤和嘧啶之间的替换称为颠换(transversion)，转换突变多于颠换突变。

碱基替换可以发生在基因组DNA序列的任何部位。当碱基替换发生在基因外DNA序列时，一般不会产生效应。如果发生在基因的调控区域，如转录因子结合的顺式作用元件，可能造成基因表达的改变。如果突变发生在基因的编码序列，导致mRNA的密码子改变，对多肽链中氨基酸序列的影响，可能出现不同的突变效应：

1. 同义突变(same sense mutation) 是指碱基替换后，一个密码子变成另一个密码子，但是所编码的氨基酸没有改变，因此并不影响蛋白质的功能。这是由于遗传密码的兼并性所致，同义突变常发生在密码子的第三个碱基，例如密码子GCA，GCG，GCC和GCU均编码丙氨酸，它们的第三碱基发生突变并不改变所编码的丙氨酸。

2. 错义突变(missense mutation) 是指碱基替换后使mRNA的密码子变成编码另一个氨基酸的密码子，改变了氨基酸序列，影响蛋白质的功能。这种突变常发生在密码子的第一或第二碱基。例如DNA序列中TCA的T突变为G，使mRNA的密码子UCA变成GCA，结果是丝氨酸被丙氨酸替换，可能使产生的蛋白质无活性或活性降低。

3. 无义突变(nonsense mutation) 是指碱基替换后，使一个编码氨基酸的密码子变为不编码任何氨基酸的终止密码子(UAG，UAA，UGA)，使多肽链的合成提前终止，肽链长度缩短，而成为无活性的多肽片段。例如正常血红蛋白β珠蛋白基因的第145密码TAT突变为TAA，mRNA上UAA为终止密码子，其结果是翻译提前终止，产生缩短的β珠蛋白链而形成了异常血红蛋白Hb Mckees Rock。

（二）移码突变

移码突变（frame shift mutation）是指在DNA编码序列中插入或丢失一个或几个碱基，造成插入点或缺失点下游的DNA编码框架全部改变，其结果是突变点以后的氨基酸序列都发生改变。例如异常血红蛋白HbW，是由于α珠蛋白基因的第138密码子TCC中的C缺失，造成该突变点以后的编码全部改变，最终的α链从第138氨基酸以后的序列不同于正常，而且没有终止于第141密码子，而是延长至第147密码子。

（三）动态突变

人类基因组中的短串联重复序列，尤其是基因编码序列或侧翼序列的三核苷酸重复，在一代代传递过程中重复次数发生明显增加，从而导致某些遗传病的发生，称为动态突变（dynamic mutation）。例如Huntington舞蹈病，是由于*huntingtin*基因5′端CAG重复序列的拷贝数增加所致。在正常人群中CAG拷贝数在9～34之间，患者拷贝数多在36～120之间。动态突变可能的机制是姐妹染色单体的不等交换或重复序列中的断裂错位。

二、DNA修复

细胞内的DNA会受到各种各样的损伤，一部分可归因于细胞外因素，但大部分是由于内源性机制，包括自发性化学水解，细胞内活性氧基团攻击嘌呤和嘧啶环以及复制与重组错误等。为了应付细胞内外这些因素对DNA的损伤，人体细胞能进行至少5种类型的DNA修复：

1. 直接修复　是通过O^6甲基鸟嘌呤DNA甲基转移酶除去不恰当的甲基化鸟嘌呤上的甲基，可以逆转DNA的损伤。

2. 碱基切除修复　即用糖苷酶除去异常碱基，可纠正许多最常见类型的DNA损伤；人体内至少有8个不同DNA糖苷酶（DNA glycoylases），每个负责识别和除去特定种类的碱基损伤。

3. 核酸酶切除修复　能除去胸腺嘧啶二聚体和大的化学加成物，该切除修复功能的缺陷可导致一种常染色体隐性遗传病——着色性干皮病。

4. 复制后修复　可纠正双链断裂，通过重组修复，使同源染色体一条单链嵌入受损的DNA，断端重接在一起。

5. 错配修复　能纠正DNA复制错误所引起的错配的碱基对。除直接修复外，所有这些系统均需要核酸外切酶及内切酶、解旋酶、聚合酶以及连接酶，它们通常以具有共同组件的多蛋白复合体形式发挥作用。首先是核酸内切酶特异识别DNA损伤部位，并在5′端作一切口，然后核酸外切酶从5′端至3′端方向切除损伤的DNA单链，同时在DNA聚合酶作用下，以损伤的互补链为模板合成新的DNA单链，最后DNA连接酶将新合成的DNA单链与原有的单链以磷酸二酯键连接而完成修复。

Summary

Genetic information is stored in DNA which is the ultimate molecule of life.DNA consists of two helical complementary strands.This structure allows DNA to find each other in a complex mixture of molecules, to defense against information loss by DNA damage, to replicate in semi-conservative mode.The genetic information flows from DNA to RNA to protein.It is the protein product of a gene that carries out its function.Almost all of the cells of the human body contain the complete genome, but only a subset of these genes are being expressed in any given tissue.Therefore, the control of gene expression is fundamental to understanding all aspects of human biology.The promoter, splicing, polyadenylation, enhancers and transcription factors are included in the control of gene expression. There are two kinds of divisions: mitosis and meiosis. Mitotic division is the process whereby one cell divides to two cells that are genetically identical to the parent.Meiosis is

used to generate the male and female gametes which bear 23 chromosomes.Mutations in DNA sequence allow evolution of organisms.They contain missense mutation, same sense mutation, nonsense mutation and frameshift mutation according to the efforts of mutations.

（赵彦艳）

参考文献

1. Turnpenny PD, Ellard S. Emery's Elements of Medical Genetics: The cellular and molecular basis of inheritance. 13th ed. Churchill: Livingstone, 2007
2. Strachan T, Read A P. 人类分子遗传学 · DNA结构和基因表达. 孙开来译. 北京:科学出版社, 2007
3. 李璞. 医学遗传学 · 遗传的细胞和分子基础. 北京:北京大学医学出版社, 2003
4. 孙开来. 人类发育与遗传学 · 配子形成与受精. 北京:科学出版社, 2008
5. Dennis C, Gallagher R. The human Genome. Nature Publishing Group, 2001

第三章 人类基因组学

基因是控制生物遗传性状的基本单位，基因组则是一个生命体遗传信息的总和。包括人类在内的绝大多数生命体的遗传信息载体是 DNA 分子。人类单倍体基因组含有 3.2×10^9 碱基对的 DNA 序列，分布于 22 条常染色体和 X、Y 性染色体。人类基因组的信息在很大程度上决定了人类发育、生长、生殖、疾病、衰老、死亡等所有生命现象。因此，阐明基因组 DNA 序列及其功能，就等于获得了一张生命的元素周期表，其科学意义不亚于阿波罗登月计划。

人类基因组计划（human genome project，HGP）是美国科学家于 1985 年在能源部（DOE）的一次会议上讨论酝酿，诺贝尔奖获得者 Renato Dulbecco 于 1986 年在《Science》杂志上发表的一篇短文中率先提出的，旨在阐明人类基因组 DNA 长达 3.2×10^9 碱基对的序列，发现所有人类基因并阐明其在染色体上的位置，从而在整体上破译人类遗传信息。美国于 1990 年正式启动 HGP，计划在 15 年内提供 30 亿美元的资助，于 2005 年完成人类基因组全部序列的测定。欧共体、日本等发达国家和巴西、印度、中国等发展中国家也相继提出了各自的基因组研究计划。由国际科学家小组建立的人类基因组组织（Human Genome Organization，HUGO）对此计划在各国的实施进行协调工作。由于各国政府、科学界和产业界的共同努力，HGP 作为全球性的合作项目已取得重大进展。

2000年6月26日，由美国、英国、日本、德国、法国、中国科学家组成的国际人类基因组测序协作组(International Human Genome Sequencing Consortium，IHGSC)和Celera公司共同宣布人类基因组草图的完成，分别在2001年2月15日的《Nature》和2月16日的《Science》上发表了他们的结果。2004年10月21日，IHGSC在《Nature》上公布了人类基因组的完成序列，该序列覆盖了约99%的常染色质区域，准确率高达99.999%，它为新世纪的生物医学研究奠定了坚实的基础。

在人类基因组计划完成之后，Collins等发表了题为"基因组学研究的未来(A Vision for Future of Genomics Research)"的文章，预见了基因组学在三个领域的重要影响：基础生物学、人类健康和人类社会，并形象地绘制出一座建立在人类基因组计划基础上的三层楼房(图3-1)。三个楼层分别称为"基因组学到生物学"、"基因组学到健康"和"基因组学到社会"。楼房的立柱包括"资源"、"技术发展"、"计算机生物学"、"培训"、"教育"和"伦理、法律及社会影响"。

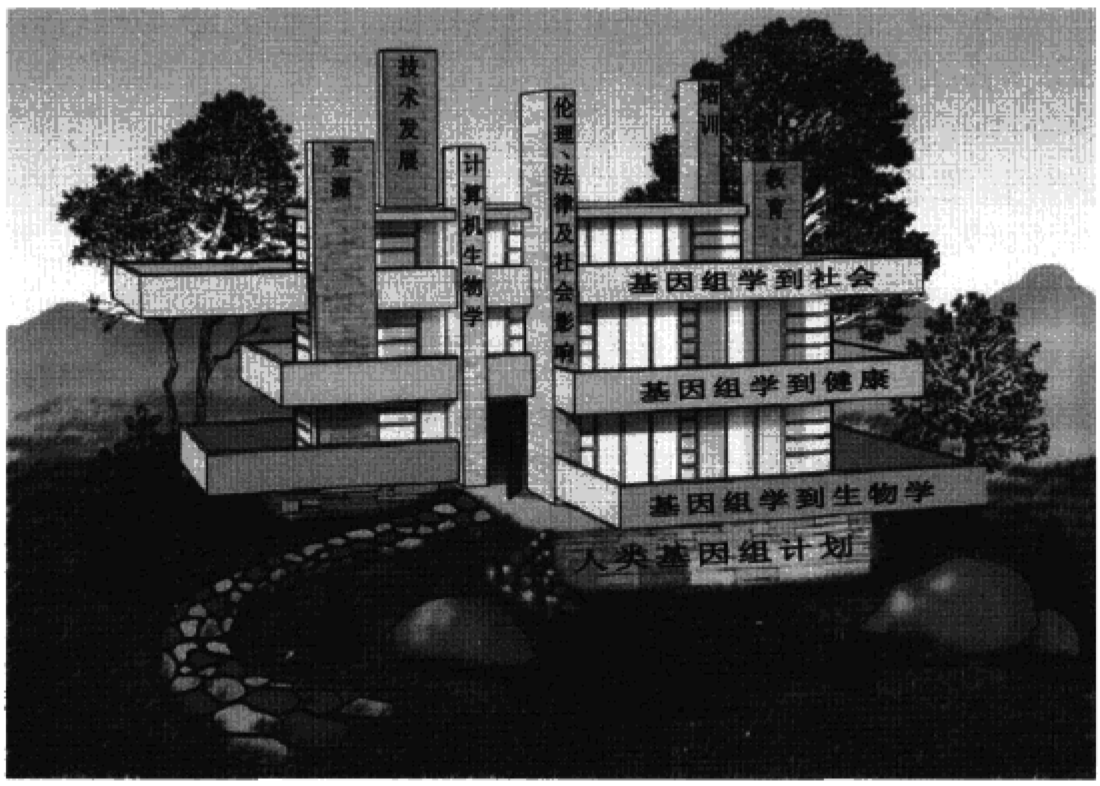


图3-1 未来的基因组学

在第一层，即基因组学到生物学，首要的任务是如何从功能上解析基因组。在这个领域，首先要寻找基因组中的功能元件(functional element)，包括基因、调控元件及其他潜在的功能区段；其次要阐明基因组表达调控的机制以及研究基因组表达产物(包括RNA和蛋白)的性质和功能；再次，要比较不同个体、不同物种之间的差异，前者可以分析人类基因组的多样性，后者则可以从进化的角度阐述基因组的功能；最后，模式生物为我们提供了个体水平研究基因功能的模型。

人类基因组学研究无疑给医学带来一场革命。目前的医学正在从"以疾病为中心"到"以健康为中心"过渡。人类基因组学对医学的贡献，不仅在于寻找病因和指导治疗，还在于维护人类健康，理解发病机制从而预防疾病的发生。另外，通过人类基因组多态性的研究，特别是大量单核苷酸多态性(single nucleotide polymorphism，SNP)的获得，可以使连锁不平衡(linkage disequilibrium，LD)分析变得简单，从而确定常见疾病的易感基因；同时，这种多态性还可以阐明个体间在药物反应上存在差异的遗传基础，从而根据个体的遗传背景优化治疗方案，实现"个体化医学"。

人类基因组计划的一小部分预算经费被用于研究"伦理、法律及社会影响(ethical，legal，and social implication，ELSI)"，人类基因组计划引发的社会问题从一开始就受到了人们的关注。

第一节 人类基因组计划

一、基因组作图

人类染色体不能直接用来进行测序,故 HGP 的第一阶段是要将基因组这一巨大的研究对象进行分解,使之成为较易操作的小的结构区域,这个过程简称为作图(mapping)。根据使用的标志和手段不同,作图可以分为两种类型:①遗传连锁图谱(genetic linkage map),即通过计算连锁的遗传标志之间的重组率确定它们的相对距离,一般用厘摩(centimorgan,cM)即每次减数分裂的重组率为1%来表示;②物理图谱(physical map),即确定各遗传标志之间的物理距离的图谱,以碱基对的个数来表示。

遗传图谱绘制需要应用多态性标志。20世纪80年代中期最早应用的标志是限制性酶切片断长度多态性(RFLP)。此类标志的数量较少,多态性信息也较低。20世纪80年代后期发展的短串联重复序列(short tandem repeat,STR)又称微卫星(microsatellite,MS)标志,主要是二核苷酸重复序列如$(CA)_n$,其染色体分布和信息含量明显优于RFLP,成为遗传连锁分析极其有用的标志。1996年法国和美国科学家共同努力,发表了由5264个MS标志组成的、分辨率高达0.6cM的遗传图谱,提前和超额完成了原定的分辨率为2~5cM的遗传作图计划。同时,MS也成为物理图谱上的标志,从而促进了遗传图谱与物理图谱的整合。近来,第三代的多态性标志,即SNP标志又被大量产生,其意义已超出了遗传作图的范围,而成为研究基因组多样性和识别、定位疾病相关基因的一种手段。

物理图谱包含了两层意义,一是获得分布于整个基因组的序列标签位点(sequence tagged site,STS),其定义为染色体定位明确,而且可用PCR扩增的单拷贝序列,使每隔一定距离就有一个标志;二是在此基础上构建覆盖每条染色体的大片断DNA(例如酵母人工染色体(yeast artificial chromosome,YAC)插入片断达数百kb至2Mb,细菌人工染色体(bacterial artificial chromosome,BAC)插入片断为80~300kb,P1噬菌体(bacteriophage P1)插入片断最大为125kb,P1来源的人工染色体(P1-derived artificial chromosome,PAC)插入片断可达300kb的重叠群(contig)。确定两个STS间的物理联系,通常用的工具是由人类家系细胞株组成的遗传作图板、人类DNA与啮齿类DNA杂交的辐射杂种(radiation hybrid,RH)细胞板和YAC/BAC。20世纪90年代中期大量产生的STS,除了部分MS以及少量非编码单拷贝序列外,主要来源是定位的表达序列标签(expressed sequence tag,EST)。EST是cDNA(complementary DNA),代表基因的转录表达产物的部分序列,经RH定位后可以作为STS。另一方面,构建覆盖人类基因组的大片断DNA重叠群也取得了如期进展。1995年法国科学家利用"指纹"法和STS构建YAC重叠群,覆盖了人类基因组75%的区域;此后,美国、欧洲和日本的科学家又利用更加稳定和易于操作的BAC构建了覆盖各条人类染色体的重叠群。物理图构建的成功,不仅为大规模测序奠定了基础,而且还绘制出了人类基因组转录图谱(或基因图谱)的雏形。

二、基因组测序

有两种基本的方法可以用于进行全基因组的序列测定。第一种是基于图谱的(map-based)的方法。通过精细物理作图,排出对应于特定染色体区域的重叠度最小的BAC重叠群后,就可以对其中的BAC逐个进行测序。这是由各国政府及非盈利机构(如英国的Welcome Trust基金)所支持的公共领域测序计划在20世纪90年代中期所确定的策略。其基本工作步骤是:①将待测BAC克隆随机切成小片断(约1.5~2kb);②将小片断克隆进测序载体;③对小片断DNA进行8~10倍左右覆盖率的测序;④将相互重叠的读出序列(reads)组装成连续的重叠线;⑤从质量最高的读出序列中取得序列;⑥利用引物延伸或其他方法对BAC克隆中还存在的缺口(gap)进行填补。由美国国立卫生研究院(NIH)和能源部(DOE)制定的最终完成序列精度标准有三条:①错误率低于1/10 000;②序列是连贯的,即没有缺口;③测序所用克隆能忠实地代表基因组结构。但在人类基因组计划进行的过程中,又

提出了工作草图(working draft)的概念,其定义为:通过对染色体位置明确的 BAC 重叠群 4~5 倍覆盖率的测序(在 BAC 克隆水平的覆盖率不应低于 3 倍),获得基因组 90% 以上的序列,其错误率应低于 1%。工作草图虽然离最终完成图还有相当大的距离,但是已具有很高的科学价值,对于基因组模体结构的认识、基因的识别、疾病基因的定位克隆、SNP 的发现、基因基本结构的解析等都非常有用。在工作草图基础上加大测序覆盖率并填补缺口,就可逐步逼近最后的完成序列:就一个 BAC 而言,其内部所有缺口被填补之后的序列成为结束序列;而对于一段染色体来说,如果 BAC 克隆群之间的缺口被填补则称为完成序列。当然,基因组中的某些部分,主要是着丝粒区域和近端着丝粒染色体的短臂部分,因存在大量重复序列而非常不稳定,是目前还难以克隆的部分。因此,基因组的最后完成序图不包括这部分的序列。

另外一种方法是"全基因组鸟枪(whole-genome shotgun)法"测序。这是一个十分大胆的构思,即在获得一定的遗传和物理图谱的信息的基础上,绕过 BAC 克隆逐个排序的过程,直接将基因组 DNA 分解成 2kb 左右的小片断进行随机测序,辅之以一定数量的 10kb 的克隆和 BAC 克隆的末端测序,利用超级计算机进行序列组装。由 Craig Venter 领导的私营研究所 TIGR 于 1995 年首先将这一策略应用于微生物基因组的测序并获得成功,但该方法能否用于高级生物的复杂基因组测序一直有疑问。1999 年,由 TIGR 和 PE 公司组建的 Celera 公司,与加州大学伯克利果蝇计划(BDGD)合作,仅用了 4 个月时间,就用该方法完成了果蝇基因组 120Mb 的全序列测定和组装,证明了这种新的技术路线的可行性,成为采用该方法进行人类基因组全序列测定的一个成功的预实验。

DNA 测序技术在 90 年代中后期出现了革命性的突破。首先,20 世纪 80 年代发展的以薄板凝胶电泳为基础的自动测序技术,实现了高通量运作。一台 PE-ABI377 测序仪每天可测 2~3 轮,每轮产生 96 条序列,每条序列的平均长度可达 500~700bp。继而,高通量的毛细管阵列电泳仪在 90 年代末问世,代表性的仪器如 MD 公司的 Megabase 和 PE-ABI 的 3700 测序仪,每天可测 10 轮,每轮 96 条序列。3700 测序仪还可以进行自动加样。近年来各国的测序中心都大量装备了这类先进仪器。高质量的聚合酶及高敏感的荧光染料相继出现,也使得测序的质量和精度不断提高。同时,数据质量评估和组装的软件系统(如 phred/phrap/consed)等不断改进和完善。为了将作图、测序和信息学技术形成高度自动化的大规模生产线并进行全过程的质量监控,实验室信息管理系统(laboratory information management system,LIMS)应运而生,该系统可以跟踪测序克隆制备的过程和从克隆中产生数据,记录试剂、方法、机器性能等参数,对于测序效率的维持和高质量数据的产生起着至关重要的作用。现在,某些特大型的测序中心或公司拥有数百台自动测序仪和大批机器人,可以每天产生数十万条读出序列,形成了生命科学领域中大科学工程的壮观景象。需要指出的是,随着微电子技术和显微制造技术的突飞猛进,今后测序技术的发展方向是仪器的微型化,这样可以极大地减少成本和提高效率。

三、基因组数据库

由 HGP 得到的图谱和序列的信息以惊人的速度积累,它只有可以被科学家免费得到才能最大限度地实现它的价值,这就需要有公共数据库来存储这些数据。1996 年 2 月,在百慕大会议上,基因组计划国际参与者对规范化数据存取方面达成一致,其中包括必须在 24 小时内把序列提交到公共数据库中,这就是所谓的"百慕大原则"。迄今,国际上三个大的生物信息中心即美国的国家生物技术信息中心(NCBI)、欧洲生物信息学研究所(EBI)和日本 DNA 数据库(DDBJ)已经建立了包含数百种生物的 DNA 序列的大型数据库。

为了使世界各地的科学工作者能够方便地浏览人类基因组的信息和抽取有价值的数据,HGP 组织还开发了各种免费供大家使用的基因组浏览器。例如位于 Santa Cruz 的加州大学开发的 UCSC 人类基因组浏览器和桑格中心(Sanger Center)创建的 Ensembl 基因组浏览器。这些基于网络的计算机工具,允许使用者浏览基因组序列和注释信息,目前的信息包括:核苷酸序列、序列重叠群、克隆重叠群、序列的覆盖度及完成的状况、局部的 GC 含量、CpG 岛、从遗传图谱和物理图谱得到的 STS 标记、

重复序列、已知基因、mRNA 和 EST、预测的基因、SNP、与其他物种基因的序列相似性。而且这些浏览器随着基因组序列的修正及注释的发展而不断更新。

四、人类基因组计划的完成

人类基因组计划于 1990 年启动,2003 年 4 月宣布完成,共耗资约 27 亿美元。其中 5% 用于研究与人类基因组有关的伦理、法律和社会相关问题,即 ELSI 计划。人类基因组计划从一开始就是一个全球合作的项目,共有来自六个国家的 20 个测序中心参与。人类基因组计划的数据全部公开,任何人都可以通过互联网免费获得整个人类基因组的序列。许多其他物种的基因组也已经被测序,包括酵母、线虫、果蝇和小鼠。还有更多物种的基因组测序工作正在进行,包括猩猩、狗、猫、牛、鸡、鱼、蜜蜂等等。数百种微生物、病毒也已经被测序,这在医学上同样重要。最近,一个可以体现基因组测序能力的事件就是迅速完成了非典(SARS)病毒的基因组测序。

人类基因组计划开始的工作重点是发展好的遗传图谱和物理图谱。在 1992 年到 2001 年之间,科学家给人类基因组构建了高精度的图谱,为大规模测序奠定了基础。同时,低等模式生物基因组的测序工作考验了测序方法并且有助于解析人类基因组序列。2001 年,IHGSC 和 Celera 公司分别公布了各自的人类基因组草图,大约 90% 的常染色质部分被测序和拼接。虽然在这两份草图中还存在大量的缺口,但是它们首次展示了人类基因组的概貌。2004 年 10 月 21 日,IHGSC 在《Nature》杂志上公布了人类基因组的完成图。与 2001 年发表的工作草图相比,基因组序列的缺口减少了 400 倍,只有 341 个。这些缺口大都位于片断复制区域,再拼接是有很大的难度,需要发展新的方法才能填补这些缺口。

五、人类基因组的特征

我们可以从人类基因组的序列中发现一些结构特征。例如:①人类基因组中含有 20 000 ~ 25 000个蛋白编码基因,不到线虫和果蝇的两倍;②许多特征在基因组上分布不均一,包括基因、转座子、GC 含量、CpG 岛和重组率等。统计结果表明它们各自的分布又有一些相关性;③大部分转座子活性明显下降,但是长散在元件 1(long interspersed element 1,LINE1)和 Alu 单元件却成为优势家族;④在染色体末端重组率较高,着丝粒附近则相反;越短的染色体重组率越高;而且重组率在不同性别之间有些差别;⑤人类基因组中含有大量的 SNP。所有的结构特征都是功能的基础,可能为我们理解人类基因组提供重要线索。

第二节　功能基因组学

人类基因组完成图的公布使我们对基因组的结构特征有了一个相对全面的认识,但是我们对基因组的功能仍然知之甚少。接下来我们面临的挑战是解析基因组的功能,即功能基因组学研究的范畴。首先我们需要寻找基因组内的功能元件并阐明其生物学意义。同时我们还需要对基因组表达调控的机制和表达产物的功能进行深入的研究。

一、基因组内的功能元件

2003 年 9 月,美国国家人类基因组研究所(National Human Genome Research Institute,NHGRI)启动了 DNA 元件百科全书(Encyclopedia of DNA Elements)计划,简称编码计划(ENCODE)。该计划的目标是鉴定人类基因组序列中所有的功能元件。在该计划的实验阶段,人类基因组的 1%(30Mb)的区段被选取进行研究。这项计划有可能会产生许多令人吃惊的结果,增进我们对基因组功能的理解。但是基于目前对于人类基因组的知识,我们已经知道的功能元件主要包括蛋白编码基因、非编码 RNA(noncoding RNA)基因、调控元件和重复序列。

（一）蛋白编码基因

确定人类基因组中的全部基因是人类基因组学研究的重要目标。随着人类基因组计划的进展，估算出的基因数量在不断改变，体现了认识的发展过程。大约在10年前，大多数科学家认为人类基因组包含了100 000个蛋白编码基因；在2001年发表的人类基因组工作草图中，科学家们认为人类基因组大概有30 000～40 000个蛋白编码基因；而人类基因组完成图的分析结果表明人类只有20 000～25 000个基因，这使我们对于人类基因组的真实情况有了更准确的了解，可以更容易地进行遗传病相关基因的定位克隆。目前，确定蛋白编码基因主要有三种方法：①cDNA序列与基因组序列相互比对可以得到基因组上的转录区域；②因为大部分基因受到选择压力而在物种间具有保守性，所以通过物种间的基因组同源性比较可以预测基因的存在；③通过计算方法预测新基因。利用不同方法得到的数据，科学家们希望建立一个完整的人类基因索引（integrated gene index，IGI）以及与之相关的蛋白索引（integrated protein index，IPI），来帮助人们深入研究人类基因和蛋白的功能。

（二）非编码RNA基因

至少有数千个人类基因以非编码RNA作为其最终产物。非编码RNA不含有开放阅读框（ORF）和多聚腺苷酸（poly A）尾巴而且较小，所以不能通过基因识别技术来发现，也很难通过cDNA文库的测序来获得（因为大部分cDNA文库是利用含有polyA尾巴的mRNA构建的）。即使我们获得完整的人类基因组序列，非编码RNA的发现仍有一定难度。虽然非编码RNA的研究领域已经产生许多令人激动的成就，但是我们还需要发展有效的方法才能系统地研究人类基因组的全部非编码RNA。

（三）调控元件

在基因组的非编码区还包含了许多顺式调控元件，如启动子、增强子等，它们参与调控大部分基因的表达。这些元件的突变会导致基因表达的失控。目前已经知道人类的某些疾病是由基因的顺式调控元件突变导致的。寻找顺式调控元件的办法主要包括实验和计算机预测的方法。最近，通过物种间基因组的同源性比较在基因组水平寻找顺式调控元件的方法显示了很高的效率，人类基因组测序工作的完成以及其他物种基因组计划的开展使得这种方法更为有效。大部分顺式调控元件是与反式作用因子相互作用来调控基因表达的，研究这种相互作用是解析基因调控网络的关键。目前已经有多种研究DNA和蛋白质相互作用的方法，例如凝胶阻滞（EMSA）和染色质免疫沉淀（Chromatin IP，ChIP）实验，特别是最近发展的基于芯片的染色质免疫沉淀（ChIP on chip）实验方法的建立和完善使得大规模研究基因调控的机制成为可能。

（四）重复序列

重复序列的数量占整个基因组的50%以上。这些重复序列常被人们称为“垃圾（junk）”，然而它们实际上是生物进化信息的重要来源。作为一个被动的标记，它们提供了研究突变和选择的方法。识别在进化过程中同时出现的重复片段并追踪它们在不同基因组区域和不同物种中的命运是完全可能的。微卫星序列被广泛用作疾病基因定位的标记就是一个很好的例子。而作为一个活跃的因素，重复序列可以通过介导异常重排而使基因组改变，产生新的基因，改变和改组现有的基因，调整全局的GC含量。它们对染色体结构和动力学研究提供了帮助，为医学遗传学和群体遗传学研究提供了有力的工具。

二、基因转录调控

在真核生物中，基因组DNA被装成染色质形式，由DNA缠绕组蛋白形成的核小体结构组成。转录前染色质结构发生一系列重要的变化，即染色质重建（chromatin remodeling）。这个过程使染色质变为一种疏松结构，从而使转录激活因子和基本转录机构能够结合。转录因子与其他蛋白形成蛋白复合体，协同性地调控靶基因的表达。这样，一个生物体可以通过启用少数调控蛋白完成多种调控作用。

基因的转录调控可以分为两种类型：①易变的调控（labile regulation）是由转录激活因子和转录

抑制因子通过在核内的浓度，共价修饰，相互作用的改变瞬时地调控基因的表达；②表观遗传学调控(epigenetic regulation)使基因表达状态相对稳定，这种状态可以从母代细胞传向子代细胞。目前已知的表观遗传学调控的机制包括染色质结构的改变和DNA、组蛋白的共价修饰。本书第十三章对表观遗传学进行了详细论述。上述两种调控类型可能会调控同一个/组基因，而它们之间的界限并不总是非常明确。

三、基因组表达产物的功能研究

虽然基因在人类基因组上只占极少的比例，但是目前认为它们包含了与生物学、医学和生物制药关系最为密切的信息。人类基因组的表达产物主要包括两大类：蛋白编码基因的产物mRNA和非编码RNA。对于前者的研究工作开始得较早，包括早期的大规模cDNA测序工作；而非编码RNA的识别仍有一定难度。我们这里主要介绍蛋白编码基因的功能研究。

（一）基因表达谱的研究

基因表达谱是指描绘特定细胞或组织在特定状态下的基因表达种类和丰度信息的数据表，它从mRNA水平上反映了细胞或组织的特异性表型和基因表达模式。基因表达谱的研究有助于理解细胞和组织的生物学功能以及阐述疾病的发生机制。同时基因表达谱还有助于发现和研究基因的功能。将各类细胞和组织的基因表达谱进行比较，就能筛出细胞特异性和发育阶段特异性的基因以及与疾病发生密切相关的基因；将不同基因表达谱的信息以染色体的位置为坐标进行整合可以得到整个基因组的转录组图谱(transcriptome map)，对于阐述基因表达调控机制将很有帮助；将不同物种的基因表达谱进行比较可以找出物种间的差异，会有助于理解基因的进化过程。研究基因表达谱的主要方法包括cDNA文库测序，微阵列(microarray)和SAGE(serial analysis of gene expression)等。

（二）全长cDNA的克隆和注释

基因的全长克隆是研究基因功能的基础。大部分的工作来源于cDNA文库的测序和EST数据的分析，目前已经获得了人类大部分基因的全长克隆。但是这种方法似乎不能覆盖所有的基因，所以最近又出现了靶向性的(targeted)全长克隆策略，主要是采用基因特异性的位于ORF两端的引物进行PCR然后克隆进统一的载体。特别是基于位点特异性重组的克隆技术(例如Invitrogen公司的Gateway克隆系统，Clontech公司的Creator克隆系统以及Stephen Elledge实验室开发的Univector克隆系统)的成熟使得这种全长克隆方法更为有效。这种方法的另外一个优点是可以将ORF直接在不同表达载体之间转换，而不用考虑酶切位点、非翻译区和框架(frame)的影响。基因全长克隆的另外一项重要工作是全长基因的注释。由于世界各地的研究者发布的全长序列没有统一的标准，所以存在一定数量的重叠。最近科学家们已经在这方面做了努力，对相当数量的全长基因进行了统一的注释，使它们更能反映人类转录组的状况。

（三）基因的分类

虽然目前的基因数据库都是按照标准格式储存数据，而且建立了强大的查询系统，但是由于基因的提交者对基因的命名和描述具有一定程度上的随意性，要想在数据库中查找某一类基因(例如转录因子)还是非常困难的。另外，在某些大规模的基因功能分析(例如基因表达谱)的工作中，研究者对基因的分类往往带有很强的主观性。所以，有必要对基因进行系统地分类并建立基因分类数据库。Gene Ontology(GO)就是一种目前已经被广泛应用的基因分类方案。GO正在建立有序的、可控制的词表(vocabulary)来描述真核生物的所有基因，对每个基因分别从三个角度(即三个最高节点)进行描述：分子功能(molecular function)、生物学过程(biological process)和细胞组分(cellular component)，每个节点下又包含子节点，例如分子功能节点包括催化活性、转录调控活性、信号转导活性等等。这样，所有的基因都可以按照特定的规则进行归类。目前，GO已经被许多基因数据库所采用。

（四）选择性剪接的研究

转录组的分析表明，选择性剪接在人类是非常普遍的。目前推测人类基因组含有20 000 ~

25 000个蛋白编码基因,但是由于选择性剪接的存在,它们将编码远远多于这个数量的蛋白。有一些选择性剪接本具有重要的生物学功能,而且也有证据说明在疾病和正常组织中许多基因的剪接模式或不同剪接本的比例会有差异,提示这些差异可能与疾病相关。人类基因组测序的完成和转录组学的发展使我们有条件系统地对这种现象进行研究,并阐明其调控机制。

(五) 蛋白组学

蛋白组学(proteomics)的定义是在生命体或细胞的整体水平研究蛋白的表达和修饰状态的学科。蛋白质是生物学活性的主要携带者,它们行使功能依赖于氨基酸顺序、三维结构、共价修饰、蛋白浓度、与其他分子的相互作用以及细胞外环境的影响。所以,蛋白质组学研究的范畴包括:①结构;②修饰;③表达水平;④定位;⑤蛋白质相互作用。与基因研究技术相比,蛋白质研究技术要更为复杂,其中包括三类主要的技术:首先,质谱(mass spectrometry)是在这个领域里应用比较广泛的一项技术。目前,二维电泳-质谱技术路线已经成为蛋白质组学的常规方法,主要用于蛋白质的识别、测序、翻译后修饰以及表达水平差异的研究。其次是基于微阵列的蛋白质组学,主要是把合成的多肽、抗原、抗体甚至细胞点在芯片上进行杂交的方法。这类方法可以用于研究蛋白(质)与蛋白(质)或其他分子(包括脂类、核酸和其他小分子)的相互作用,以及翻译后修饰的研究,同时它在药物发现和临床诊断上也有很好的应用前景。第三是将结构生物学的方法应用于蛋白质组学,包括X射线晶体衍射、核磁共振等技术,主要用于解析蛋白复合体的结构以及不同亚基之间的关系。除了这几类关键技术以外,蛋白质组学的一个特点就是更加依赖生物信息学。蛋白质组学产生的大量数据需要计算机来处理,得到的结果要以数据库的方式进行存储。蛋白质组学数据库范围很广,包括蛋白序列数据库、质谱数据库、二维电泳图谱数据库、结构数据库以及相互作用数据库等等。最近,在人类蛋白质组组织(human proteome organization,HUPO)的协调下,人类蛋白质组计划(human proteome project,HPP)已经启动。

四、基因组多样性的研究

人类是一个具有多态性的群体。不同群体和个体在生物学性状以及在对疾病的易感性/抗性上的差别,反映了进化过程中基因组与内、外环境相互作用的结果。开展人类基因组多样性的系统研究,无论对于了解人类的起源、进化和迁徙,还是对于生物医学均会产生重大影响。已知人类基因组DNA序列中最常见的变异形式是SNP,在全基因组中估计有$(3\sim10)\times10^6$个。与罕见的单核苷酸变异所不同的是,SNP等位基因的频率应不低于1%。当SNP位于基因的编码序列中即称为cSNP。若cSNP引起蛋白质重要部位氨基酸的变异,可导致其功能改变;位于基因调控序列中的SNP则可能影响基因表达的剂量。所以这两种SNP的生物学意义更为显著,是基因组中决定人类表型多样性的核心信息。另一方面,SNP因连锁不平衡所形成的单倍型也可用于关联研究(association studies)来确定与之连锁的生物学性状相关序列。目前,已发展了多种自动化和批量化检测SNP的技术,其应用范围十分广泛,包括连锁分析与基因定位、疾病的关联研究、多基因疾病的基因定位、个体识别和亲子鉴定,发病机制的研究,以及研究生物进化、物种间相互关系等。2002年10月启动的国际人类基因组单倍型图谱(HapMap)计划,目标是构建整个人类基因组DNA序列中多态位点的常见模式,对至少100万SNP进行全基因组规模的基因分型检测。该计划将由加拿大、中国、日本、英国和美国的10个研究中心进行。

值得指出,目前已发现的大多数SNP属于全球人群中随机频率较高的变异,因而也就是人类进化早期阶段(约10~20万年前"走出非洲"的时刻)的"老"的SNP(约占全部SNP的85%),这些SNP的LD程度较低。已知不同人群间的SNP频率可以有相当大的差别,某些SNP甚至呈现群体转移性(两者相加约占全部SNP的15%),选择这些更为"年轻"的,具有群体特异性的SNP,可能更适合基于LD的关联分析。

在基因组多样性研究方面,一个近年备受医学界和制药工业界关注的新领域是药物基因组学

(pharmacogenomics)。药物的疗效和副作用受到机体多种因素的影响,尤其是药物代谢酶、转运体、受体和其他药物靶点蛋白,而编码这些蛋白的基因在不同个体间存在着遗传多样性,其基本形式也是SNP。药物基因组学就是要阐明个体间在药物代谢和效应方面发生差别的遗传基础,促进新药的发现,并根据个体的遗传背景来优化药物治疗方案,亦即"个体化治疗"。这一研究同时也能使某些药物找到合适的治疗人群。本书第十章将深入论述药物基因组学的基本原理和具体应用。

五、比较基因组学

比较基因组学(comparative genomics)是对不同物种的基因组及其表达产物进行比较研究的学科。它是对基因组进行功能注释的主要方法之一。基因组内的功能区域处于进化选择的压力之下,被标上了特殊的"记号"。例如,阴性选择(negative selection)使它们的进化速度低于平均水平;阳性选择(positive selection)使它们的进化速度高于平均水平。物种间比较基因组序列可以找到这些记号从而推测这些区域是有功能的。这种策略在编码(ENCODE)计划中受到重视。目前正在研究的1%的区段已经经过了初步的比较基因组学分析,其结果放在公共数据库内供世界各地的研究者参考。

比较基因预测(comparative gene prediction)是应用比较基因组学的一个例子。如前所述,通过比较基因组学进行基因预测是一种重要的基因预测方法。在进化关系比较接近的物种(例如人和小鼠)之间,蛋白编码基因的外显子的变化速率远远低于周围的内含子和其他非编码区域,同时基因的结构也非常相似。这些规律使我们可以从基因组上确定基因的位置。相似的方法也被用来寻找顺式调控元件。很多启动子、增强子具有较高的保守性。目前已经有许多报道应用大规模的基因组比较来寻找和分析这些调控元件。其他的调控元件,如沉默子(silencer)、隔离子(insulator)或边界元件(boundary element)缺乏明确的特征,所以很少在保守区域中被发现。

在人类基因组中,大部分基因(外显子)受到阴性选择的压力而变化速度较慢(实际上,受到阴性选择影响的不仅包括基因,还包括其他功能元件;据估计,至少有5%的人类基因组受到阴性选择的压力,远远高于外显子的总量)。但是,仍然有少数的基因受到阳性选择的影响,它们进化的速度要快于基因组的平均水平。这些基因让人更感兴趣,因为它们可能在一定程度上代表了人类区别于其他哺乳动物的特征。IHGSC通过对人类基因组完成图进行的比较基因组学分析找到了1183个这样的基因,它们在小鼠中没有直向同源物。有趣的是,这些基因中的相当一部分与免疫、嗅觉和生殖功能有关,提示人类在这些方面可能与啮齿类动物有较大差异。作者又分析了32个在小鼠中依然有活性的人类的假基因,又发现其中10个是嗅觉受体。所以在人类进化过程中嗅觉受体基因产生得多,退化得也多。但是人类的嗅觉功能似乎仍然不如小鼠,可能是因为人类有功能的嗅觉受体的数量比不上小鼠。

六、模式生物体

如前所述,在人类基因组的研究中,模式生物体(model organism)的研究占有极其重要的地位。通过比较和鉴别进化不同阶段生物体的基因组信息,将进一步加深对人类基因组结构和功能的了解。模式生物体的基因组结构相对简单,但是它们的核心细胞过程和生化通路在很大程度上是保守的。同时,人类的发育阶段在一定程度上可以看做进化过程的重演,即"进化-发育(Evol-Dev)理论"。所以不同进化阶段的模式生物可以在不同程度上模拟人类发育过程,用于研究人类的发育。而且,人们已经利用模式生物建立了多种人类疾病模型,极大地促进了生物医学的发展。

识别基因功能最有效的方法可能是观察基因表达被阻断后在细胞和整体所产生的表型变化。在这方面,基因剔除(knockout)模式生物体成为特别有用的工具。目前,国际上已实现了对酵母所有基因的缺失突变体的构建。随着线虫和果蝇基因组测序的完成,也可以对这两种生物展开类似的研究。小鼠作为哺乳类中代表性的模式生物,在功能基因组学研究中占有特殊的地位。近年来发展的条件性基因剔除技术,已可达到对任何基因在不同发育阶段和不同器官、组织的选择性剔除。除了用同源

重组技术制造基因剔除模型,也可用化学诱变剂或插入突变方法随机形成模式生物体的基因突变,对产生表型变化者利用快速基因定位法识别致病基因。例如,德国科学家率先应用突变诱导剂对小鼠胚胎干细胞进行大规模随机诱变和表型筛查,取得了很大的成功。此外,近来也有人利用组合化学方法尝试发展针对蛋白质的化学"剔除"试剂,用来激活或失活各种蛋白质。值得一提的是,斑马鱼(zebrafish)作为一种新的模式生物,最近已经广泛地应用于功能基因组学研究。它有许多优点,如个体小、生长周期短、产卵多、体外受精、胚胎透明,所以特别适用于大规模的突变筛选。斑马鱼作为一种脊椎动物,许多生理过程与人类相似,可以用来研究人类的发育、疾病、衰老等过程。另外,它还是药物筛选的理想模型。

第三节　基因组学与人类健康

人类基因组计划的直接始动因素是要解决包括肿瘤在内的人类疾病的分子遗传学问题,因此与人类健康密切相关。另一方面,6000 多种单基因遗传病和多种大面积危害人群健康的多基因疾病(如肿瘤、心血管病、代谢性疾病、神经-精神类疾病、免疫性疾病等)的致病基因和相关基因,代表了具有一定生物学冗余性的所有人类基因中结构和功能完整性至关重要的那一部分。因此,疾病基因的定位、克隆和鉴定,是人类基因组计划的各种竞争中居于核心的部分,也是 HGP 启动以来在社会上显示度最大的成就。疾病基因的发现,还直接促进了基因诊断和靶向治疗的发展。

一、疾病相关基因的识别

人类基因组序列的最直接的应用就是促进了疾病相关基因的识别。基因组在结构、表达上的异常可能是人类大部分疾病的根本原因。所以,疾病相关基因的识别也包括两个层次:结构和表达。同时,疾病相关基因网络的概念,提供了深入了解疾病机理的新思路。

确定疾病相关基因最常规的方法是遗传连锁分析,通常是用大量的 DNA 多态标志对受累家系进行连锁分析初步定位与疾病相关的染色体区段,然后克隆出该区段的 DNA 片断,再从中寻找突变,这就是定位克隆方法。这种方法在人类基因组计划的遗传和物理作图技术的推动下迅速发展起来,导致了包括囊性纤维变性、Huntington 舞蹈病、遗传性结肠癌、乳腺癌等一大批重要疾病基因的发现(可参考本书第五章单基因遗传病)。随着人类基因组计划的进展,特别是最近完成图的公布,所有人类基因很快就会被精确地定位于染色体的各个区域。因此,一旦某个疾病位点被定位,即可从局部的序列中遴选出结构、功能相关的基因进行分析,这就是"定位候选克隆"的策略。

当前,多基因病已成为疾病基因组学研究的重点,多基因病不遵循孟德尔遗传表型-基因型的规律,难以用一般的家系遗传连锁分析取得突破。过去数年中,已发展了受累同胞对分析、关联分析和连锁不平衡分析、基于家系背景的连锁不平衡分析等多种方法,结合 MS 和 SNP 等高度多态性标志的应用,对家系和人群进行疾病相关位点在基因组中的定位(多基因疾病的基本原理和分析方法在本书第六章有详细论述)。一般认为,隔离人群或较均一的人群可能是进行多基因性状定位研究较理想的群体。最近启动的单倍型图谱计划将有效地促进多基因疾病的研究。

动物模型为多基因疾病的分析提供了有效的工具。由于可以进行选择性交叉育种,对于某些模式生物的易感基因进行精确定位常常更加容易。值得注意的是,在其他物种中检出的易感基因并不一定直接与人类相关。但是,如果模式生物基因组某一区域中的易感基因与人类基因组上的同源基因在染色体位置上相关,即存在同线性(synteny),这一模型将非常有意义。这种同线性的分析是应用模式生物进行人类遗传病研究的有效工具。

另外,疾病相关基因的概念不应仅仅包括已知的蛋白编码基因;基因组中其他的功能元件也与疾病有关。特别是许多顺式调控元件控制着基因的时间和空间的表达模式,目前已经有许多证据说明这些调控元件在疾病发生过程中的重要角色。生长因子受体 Her2/Neu 在一种类型的乳腺癌中的过

表达就是一个很好的例子。我们目前对基因组中的功能元件知之甚少,但最近启动的编码(ENCODE)计划致力于识别人类基因组中所有的功能元件,将为我们寻找疾病相关基因提供新的研究对象。

另一类寻找疾病相关基因的方法不是基于基因序列的分析,而是通过研究疾病状态下基因表达的变化。首先,基因表达调控的异常可能是导致疾病的重要因素。其中表观遗传学调控与人类疾病的密切关系在前面已有描述。一些大规模检测表观遗传型的方法,特别是检测 DNA 甲基化的方法已经成熟;人类表观基因组计划(HEP)将识别人类基因组中所有的甲基化可变位点。这些进展将使我们可以从表观遗传学水平寻找疾病相关基因。其次,一个基因编码的 mRNA 或蛋白在疾病和正常组织的表达有显著的差异,可以作为该基因与疾病相关的间接证据。DNA 阵列技术可以同时检测上万个基因在疾病和正常组织之间的表达差异。应用这种技术,已经发现了许多复杂疾病(如癌症、糖尿病、精神分裂症等)的相关基因。人类基因组计划的完成使得芯片上的基因数目更多,分类更细致;许多公司目前正在出售涵盖了所有人类已知基因的芯片。

二、基于基因组的诊断学

确定疾病相关基因的序列和表达模式,将极大地促进遗传病的诊断技术。据估计,大约有 3000 种人类单基因疾病,其中一半以上的疾病基因已经被识别,而且新发现的疾病基因正在以大约每周 5 个的速度增长。这些基因与疾病的关系一旦被确定,它们将很快被用于疾病诊断,虽然有一些基因专利限制了它们的使用。迅速增加的基因诊断方法使得许多疾病得以精确诊断。

许多常见病,如心血管病、糖尿病、肿瘤等,属于多基因病,其患病风险有一部分取决于遗传学因素,另外一部分则取决于环境因素,例如吸烟、饮食和锻炼。而且它们的遗传学因素来自于多个基因的共同作用,每个基因对于患病风险的贡献不足 5%,所以很难用单个基因进行诊断。在这种情况下,基因表达可以被用于疾病的诊断。广泛应用的工具是 DNA 微阵列。它在早期曾被用于淋巴瘤和乳腺癌的分型,现在已经广泛应用于许多其他肿瘤,包括脑癌、卵巢癌、肺癌、结肠癌、肾癌、前列腺癌、胃癌和白血病。最近,首次大型的基于基因表达谱的乳腺癌治疗项目已经展开。另外,蛋白质组的研究,特别是最近开展的人类蛋白质组计划(HPP)可能会提供在蛋白质水平上进行疾病诊断的方法。

由于遗传学因素在疾病诊断和治疗中的重要性,所有的医师都需要熟悉遗传学原理和遗传学检测,才能正确地诊断、治疗疾病以及为患者提供建议。医学遗传学,曾经作为专家的领域,现在很快变成了初级医师的必修课。因此,美国国家人类基因组研究所(NHGRI)和健康资源和服务管理局(Health Resources and Services Administration)最近资助了一项名为"初级护理中的遗传学(Genetics in Primary Care)"的课程,来提高初级医师的遗传学水平。

三、基于基因组的靶向治疗

以疾病相关基因或蛋白作为药物靶点,是开发新药的有效手段。了解人类基因组,将会大大地扩展药物靶点的范围。人类基因组包含 20 000 ~ 25 000 个基因,而蛋白的数量要远多于此,就算只有很少一部分蛋白可以作为药物靶点,预期它们的数目会有数千个。然而目前,只有不到 500 个蛋白成为小分子药物的靶点,也就是说,还有大约 90% 的蛋白可以作为潜在的药物靶点。当然,除了蛋白以外,各种 RNA 以及基因组内的调控元件都有可能成为药物的靶点。目前,已经具备了可以大幅度加快药物开发的速度的几个条件:①人类基因组计划以及功能基因组学研究提供了大量的潜在的药物靶点;②合成化学技术提供了高质量的化合物文库;③自动化技术和信息学的发展使得大规模、高通量的药物筛选成为可能。

人类基因组对靶向治疗的另外一个贡献是可以产生更好的药物。首先,基于疾病机理的治疗策略比原始的基于表型的治疗方法具有更好的效果。几个用于治疗肿瘤的药物就是很好的例子:Trastuzumab(Herceptin),一个针对生长因子受体 Her2/Neu 的单抗,用于治疗乳腺癌;imatinib(Gleevec),

Bcr-Abl 酪氨酸激酶抑制剂，用于治疗慢性粒细胞性白血病（CML）；bortezomib（Velcade），一种蛋白酶体的抑制剂，可以抑制多种骨髓瘤中的 NFκB 途径的激活。这几种药物在临床上都显示了明显的疗效。其次，基因阵列技术用于确定新药的特异性和作用机制，可以更为精确地评价药物在分子水平的效果。FDA 已经联合制药和生物技术公司开始制定阵列数据发布的标准，目的是加快这些数据在药物开发上的应用。

四、基因组与环境的相互作用

任何生物体都是一个开放的系统，只有与外界环境发生物质和能量的交流，才能维持自身高度有序的状态，这是基于热力学第二定律的系统生物学的观点。基于这种观点，我们很容易认识到，绝大多数人类疾病是基因组信息与环境因子相互作用的结果。1997 年美国提出了环境基因组学计划（environmental genome project，EGP），其目的是要了解环境对人类疾病的影响和意义。由于人类遗传的多态性，不同个体对环境致病因素的易感性也有差异。针对与环境中物理、化学或生物因素发生相互作用的蛋白的编码基因（如 DNA 修复机制、氧化-还原反应及病毒受体蛋白等），识别其基因组多样性和结构-功能关系，将有助于发现特定环境因子致病的风险人群，并制定相应的预防措施和环境保护策略。

Summary

The human genome is found in nuclear DNA and mitochondrial DNA. In order to understand all the sequences and genes of human beings, the Department of Energy(DOE), Human Genome Program and the NIH National Human Genome Research Institute (NHGRI) in USA began the Human Genome Project(HGP) in 1990. Then, the International Human Genome Sequencing Consortium (IHGSC) completed this project in 2003.

According to the annotation of the finished DNA sequence of human genome by IHGSC, there are about 25 000 protein-coding genes in the human genome, accounting for 1.1% ~ 1.4% of the entire human genomic sequence. However, it has been found that human genome possesses a huge quantity of single nucleotide polymorphism (SNP), about $(3\sim10)\times10^6$ in number, which was revealed over the course of the HGP.

Understanding of the biological meanings of the non-coding sequences, which represent a large majority of the genomic sequence, is a question faced by the medical genetics community in the 21st century.

（陈 竺）

参 考 文 献

1. 陈赛娟. 人类基因组研究基本技术. 北京：人民军医出版社，2002
2. 陈竺. 医学遗传学. 北京：人民卫生出版社，2001
3. 陈竺，强伯勤，方福德. 基因组科学与人类疾病. 北京：科学出版社，2001
4. C. 丹尼斯，R. 加拉格尔. 林侠 等译校. 人类基因组——我们的 DNA. 北京：科学出版社，2003
5. T. D. 盖莱哈特，F. S. 柯林斯，D. 金斯伯格. 孙开来 等译. 医学遗传学原理. 北京：科学出版社，2001
6. Austin CP. The impact of the completed human genome sequence on the development of novel therapeutics for human disease. Annu Rev Med, 2004, 55: 1-13
7. Chu JY, Huang W, Kuang SQ, et al. Genetic relationship of populations in China. Proc Natl Acad Sci U S A, 1998, 95:

11763-11768

8. Collins FS, Green ED, Guttmacher AE, et al. A vision for the future of genomics research. Nature, 2003, 422: 835-847
9. Collins FS, Patrinos A, Jordan E, et al. New goals for the U. S. Human Genome Project: 1998 ~ 2003. Science, 1998, 282: 682-689
10. de Hoog CL, Mann M. Proteomics. Annu Rev Genomics Hum Genet, 2004, 5: 267-293
11. Dulbecco R. A turning point in cancer research: sequencing the human genome. Science, 1986, 231: 1055-1056
12. Hu RM, Han ZG, Song HD, et al. Gene expression profiling in the human hypothalamus-pituitary-adrenal axis and full-length cDNA cloning. Proc Natl Acad Sci U S A, 2000, 97: 9543-9748
13. Imanishi T, Itoh T, Suzuki Y, et al. Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS Biol, 2004, 2: e162
14. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature, 2004, 431: 931-945
15. Jiang YH, Bressler J, Beaudet AL. Epigenetics and human disease. Annu Rev Genomics Hum Genet, 2004, 5: 479-510
16. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature, 2001, 409: 860-921
17. Liu TX, Zhou Y, Kanki JP, et al. Evolutionary conservation of zebrafish linkage group 14 with frequently deleted regions of human chromosome 5 in myeloid malignancies. Proc Natl Acad Sci U S A, 2002, 99: 6136-6141
18. Marsischky G, LaBaer J. Many paths to many clones: a comparative look at high-throughput cloning methods. Genome Res, 2004, 14: 2020-2028
19. Miller W, Makova KD, Nekrutenko A, et al. Comparative genomics. Annu Rev Genomics Hum Genet, 2004, 5: 15-56
20. Song HD, Sun XJ, Deng M, et al. Hematopoietic gene expression profile in zebrafish kidney marrow. Proc Natl Acad Sci U S A, 2004, 101: 16240-16245
21. Swinbanks D. Genome research set to take off in China. Nature, 1998, 394: 109

第四章　人类染色体与染色体病

染色体(chromosome)是遗传物质—基因的载体。真核细胞的基因大部分存在于细胞核内的染色体上,通过细胞分裂,基因随着染色体的传递而传递。在不同物种中,染色体的数目、形态结构、大小各具特征;而在同一物种中,染色体的形态结构、数目是恒定的。所以染色体如果发生了异常,无论是数目还是结构的畸变,都会导致许多基因的增加或缺失。因此,染色体异常常表现为具有多种畸形的综合征,又称为染色体综合征。

第一节　人类染色体的基本特征

一、染色质与染色体

染色质(chromatin)与染色体(chromosome)是同一种物质在细胞周期的不同时期中所表现的不同存在形式。染色质和染色体是一种由 DNA、组蛋白、非组蛋白及 RNA 等组成的核蛋白复合物,是核基因的载体。染色质是细胞间期核内伸展的 DNA 蛋白质纤维,而染色体则是高度螺旋化的 DNA 蛋白质纤维,是间期染色质结构紧密盘绕折叠的结果。

(一) 染色质

染色质是 DNA 和蛋白质的复合体。伸展的染色质在电镜下呈现出串珠样的结构,珠间由细丝连起来,每一珠体与其旁的珠间细丝为一个单元,名为核小体。其中珠体为核小体的核心,珠间细丝为连接区。因此,染色质是由一条 DNA 分子缠绕无数核小体核心组成的核蛋白纤维。间期细胞核的染色质可根据其所含核蛋白分子螺旋化程度以及功能状态的不同,分为常染色质(euchromatin)和异染色质(heterochromatin)两类。

常染色质在细胞间期螺旋化程度低，呈松散状，染色较浅而均匀，含有单一或重复序列的 DNA，具有转录活性，常位于间期细胞核的中央部位。异染色质在细胞间期螺旋化程度较高，呈凝集状态，而且染色较深，多分布在核膜内表面，其 DNA 复制较晚，含有重复 DNA 序列，很少进行转录或无转录活性，为间期核中不活跃的染色质。异染色质又分为两种：一种称为专性异染色质或称结构异染色质（constitutive heterochromatin），结构异染色质是异染色质的主要类型，这类异染色质在各种细胞中总是处于凝缩状态，一般为高度重复的 DNA 序列，没有转录活性，常见于染色体的着丝粒区、端粒区、次级缢痕，以及 Y 染色体长臂远端 2/3 区段等。另一种为兼性异染色质（facultative heterochromatin），也叫功能异染色质，这类染色质是在特定细胞或在某一特定发育阶段，由常染色质凝缩转变而形成的。在浓缩时基因失去了活性，无转录功能，当其处于松散状态时，又能够转变为常染色质，恢复其转录活性。如 X 染色质就是一种兼性异染色质。

（二）染色体

染色体是由染色质通过多级螺旋包装形成。每条染色体在复制前含有一条 DNA 双螺旋分子。人类的一个基因组 DNA 含有约 3.2×10^9 碱基对，平均每一条染色体的 DNA 含有 1.3×10^8 碱基对，以每一碱基对间相距 0.34nm 来计算，每一条染色体上的 DNA 的总长度约有 5cm。染色质的基本单位是核小体（nucleosome）。核小体由核心颗粒（core particle）和连接区（linker）两部分组成。核心颗粒的核心是由四种组蛋白（H2A，H2B，H3，H4 各 2 个分子）所形成的八聚体以及围绕在八聚体周围的 DNA 所组成。其直径约 11nm。这段 DNA 称为核心 DNA，约有 146 个碱基对，围绕核心颗粒外周 $1\frac{3}{4}$圈。两个核心颗粒之间的 DNA 链称为连接区，这段 DNA 长约 60 个碱基对。组蛋白 H1 位于连接区 DNA 表面。连接区 DNA 的长度差异较大，短的只有 8bp，长的可达 114bp。无数个重复的亚单位—核小体通过一条 DNA 分子串联起来，形成一条串珠状的纤维。这就是染色体的一级结构，DNA 长度被压缩了 7 倍。由核小体构成的串珠状纤维进一步螺旋化，形成螺线管（solenoid），DNA 的长度又被压缩了 6 倍，螺线管是染色体的二级结构。由螺线管进一步的螺旋化形成超螺线管（super solenoid），此时 DNA 的长度又被压缩了 40 倍，超螺线管是染色体的三级结构。由超螺线管再缠绕折叠形成了有丝分裂中期的染色体，DNA 又压缩了 5 倍，是染色体的四级结构。这样经过几级包装，染色体中的 DNA 长度就被压缩了近万倍（图 4-1）。

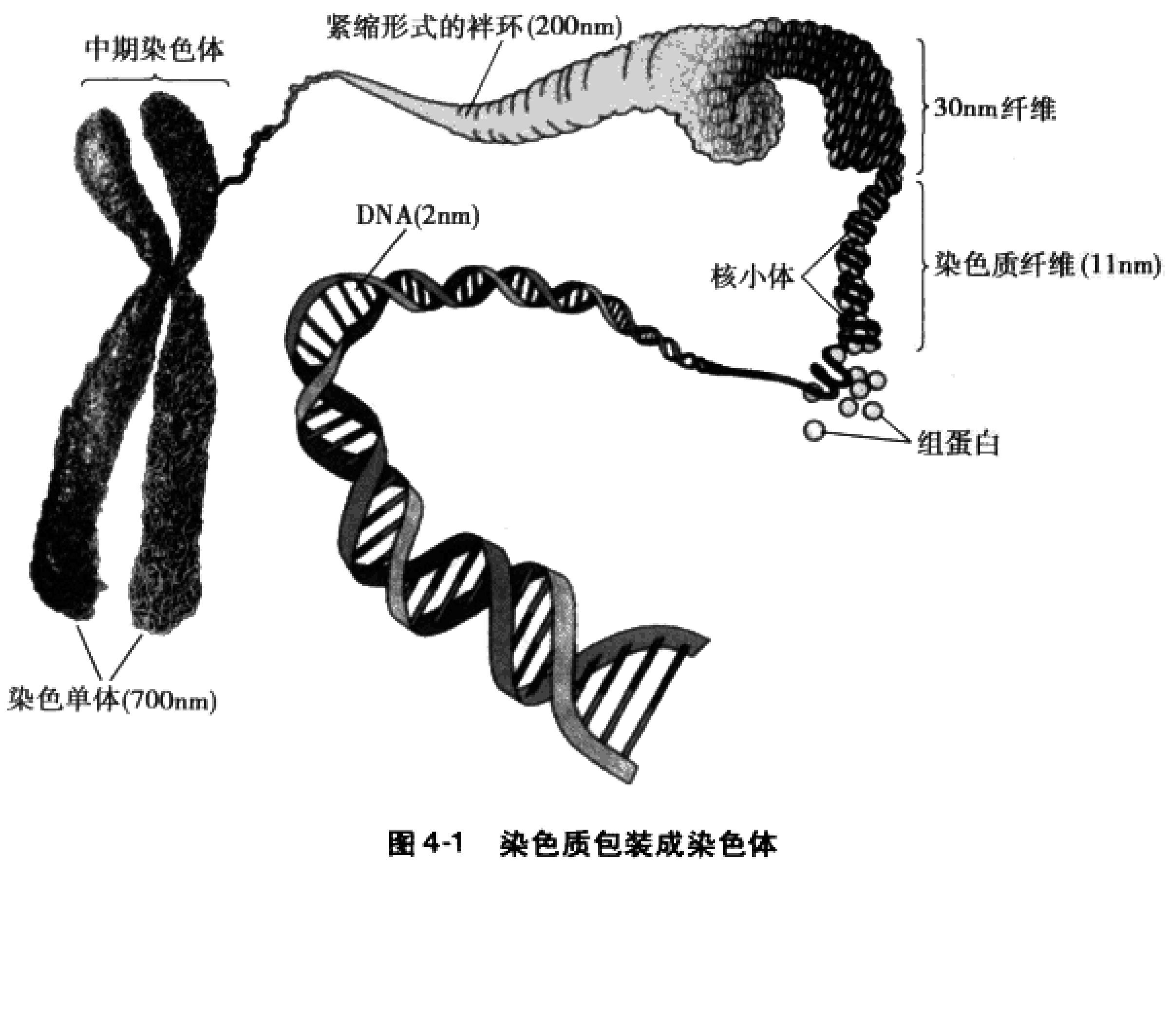


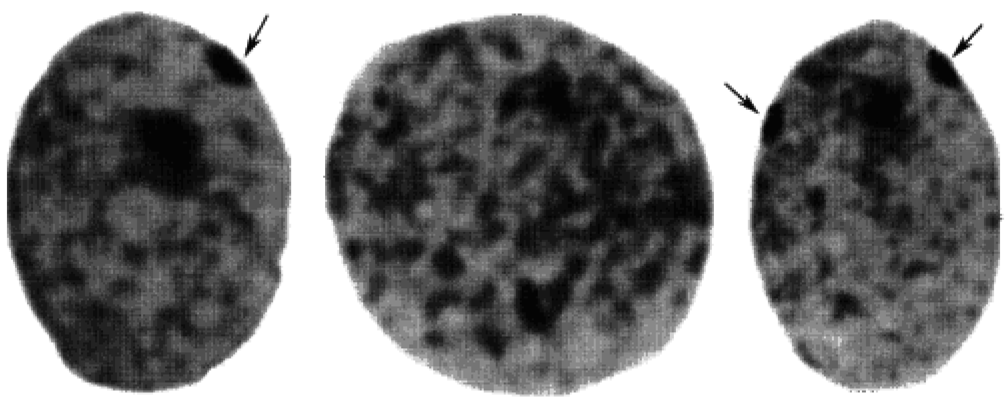
图 4-1 染色质包装成染色体

染色体的形态结构在细胞增殖周期中是不断运动变化的，一般在细胞分裂中期染色体的形态是最典型的，很容易辨认和区别，可以在光学显微镜下观察。

（三）性染色质

性染色质（sex chromatin）是在间期细胞核中性染色体的异染色质部分显示出来的一种特殊结构。人类性染色体有 X 和 Y 两种，所以性染色质也有 X 染色质（X-chromatin）和 Y 染色质（Y-chromatin）。

1. X 染色质　1949 年 Barr 等人在雌猫神经元细胞核中观察到一个浓缩小体，在雄猫中则见不到这一结构。进一步研究发现，除了猫以外，其他雌性哺乳类动物（包括人类）也同样有这种显示性别差异的结构。而且不仅是神经元细胞，在其他细胞的间期核中也可以观察到这一结构，称之为 X 染色质，也称 Barr 小体或 X 小体（图 4-2）。



1个X染色质　　无X染色质　　2个X染色质

图 4-2　间期核 X 染色质

箭头所指为染色质

正常女性的间期细胞核中紧贴核膜内缘有一个染色较深，大小约为 1μm 的椭圆形小体，即 X 染色质。正常男性则没有 X 染色质。为什么正常男女性之间的 X 染色质存在差异？女性两个 X 染色体上的每个基因座的两个等位基因所形成的产物，为什么不比只有一个 X 染色体半合子的男性的相应基因产物多？为什么某一 X 连锁的突变基因纯合子女性的病情并不比半合子的男性严重？1961 年，Mary Lyon 提出了 X 染色体失活的假说，即 Lyon 假说，对这些问题进行了解释。这一假说提出的实验根据是对小鼠 X 连锁的毛色基因的遗传学观察。发现雌性小鼠毛色的杂合体不表现显性性状，也不是中间类型，而是显性和隐性两种颜色嵌合组成斑点状（不是共显性）。而雄性小鼠却从不表现斑点状毛色，而是显性或隐性单一的毛色。Lyon 同时也注意到间期核内 X 染色质数目总是比 X 染色体数目少 1，即 XX 者有 1 个 X 染色质，XXX 者有 2 个 X 染色质。因此，两个 X 染色体中有 1 个 X 染色体是异固缩的，并且是迟复制的。在细胞代谢中，异固缩的 X 染色体没有活性，只有 1 个 X 染色体有活性。在异常细胞中具有的额外 X 染色体也无活性。对于正常男性，单个的 X 染色体不发生异固缩，而且任何时候都是有活性的，故无 X 染色质。

Lyon 假说的要点如下：

（1）失活发生在胚胎发育早期。X 染色体随机失活发生在女性早期发育阶段。女性胚胎形成之后很短一个时期，遗传自父方和母方的 X 染色体均有活性。胚胎发育的第一周，通过位于 Xq13.2 的 X 失活中心等，使某条 X 发生随机失活。

（2）X 染色体的失活是随机的。异固缩的 X 染色体可以来自父亲也可以来自母亲。

（3）失活是完全的。雌性哺乳动物体细胞内仅有一条 X 染色体是有活性的。另一条 X 染色体在遗传上是失活的，在间期细胞核中螺旋化而呈异固缩为 X 染色质。

（4）失活是永久的和克隆式繁殖的。一旦某一特定的细胞内的 X 染色体失活，那么由该细胞而增殖的所有子代细胞也总是这一个 X 染色体失活。即原来是父源的 X 染色体失活，则其子细胞中失

活的 X 染色体也是父源的，也就是说，所有这个细胞的子代细胞中都将表达有活性的母源 X 染色体，而父源 X 染色体失活。因此，失活是随机的，但是，是恒定的。因此，就 X 连锁基因的表达来说，女性为嵌合体；一些细胞只表达从父方遗传来的 X 染色体等位基因，而另一些细胞则恰好相反。

研究表明，当细胞内 X 染色体数目超过两条时，仍只有一条保持活性，其余的都形成异固缩的 X 染色质。正常男性只有一条 X 染色体，所以 X 染色质数目为零。45,X 性腺发育不全的患者虽然是女性，但是因为只有一条 X 染色体，所以细胞内无 X 染色质。一个细胞中所含的 X 染色质的数目等于 X 染色体数目减一。由于雌性细胞中的两个 X 染色体中的一个发生异固缩，也称为 Lyon 化现象，失去转录活性，这样保证了雌雄两性细胞中都只有一条 X 染色体保持转录活性，使两性 X 连锁基因产物的量保持在相同水平上。这种效应称为 X 染色体的剂量补偿（dosage compensation）。

失活 X 染色体上有一个 X 失活中心，是在 680～1200kb 的区段内，导致 X 染色体特异性失活的位点。通过对结构异常的研究，已经将失活 X 染色体的 X 失活中心定位于 Xq13 条带上。X 失活中心包含一个特殊的基因 *XIST*（inactive specific transcripts），可能是 X 失活的一个关键的调控基因。*XIST* 的显著特点是只有位于失活 X 上的等位基因表达；而在男性或女性细胞的活性 X 上都不进行转录。虽然未知 *XIST* 的具体作用机制，但缺失 *XIST* 则不会发生 X 失活。*XIST* 的产物是细胞核中与失活 X 染色体紧密关联的非编码 RNA。

需要指出的是，虽然 X 染色体失活通常是随机的，是染色体的一个特征，但不是完全的，并非失活的 X 染色体上所有基因都失去了活性，有一部分基因仍保持一定活性。对 X 连锁基因的深入研究表明，至少有 15% 的基因逃脱了这种失活（图 4-3），无论在活性 X 染色体和失活 X 染色体上都可以表达。因此，X 染色体数目异常的个体在表型上有别于正常个体，出现多种异常临床症状。如 47,XXY 的个体不同于 46,XY 的个体；47,XXX 的个体不同于 46,XX 的个体，而且 X 染色体数目越多时，表型的异常更严重。另外 10% 的基因显示了失活的多样化；在某些女性中表达，而在另一些女性中不表达。显然，这些基因不是随机分布于 X 染色体上的；逃脱失活的基因大多（约有 50%）位于 Xp 远端，少部分位于 Xq（见图 4-3）。这对于部分 X 染色体非整倍体的遗传咨询有重要意义，因为 Xp 上基因的不平衡比 Xq 可能产生更大的临床影响。

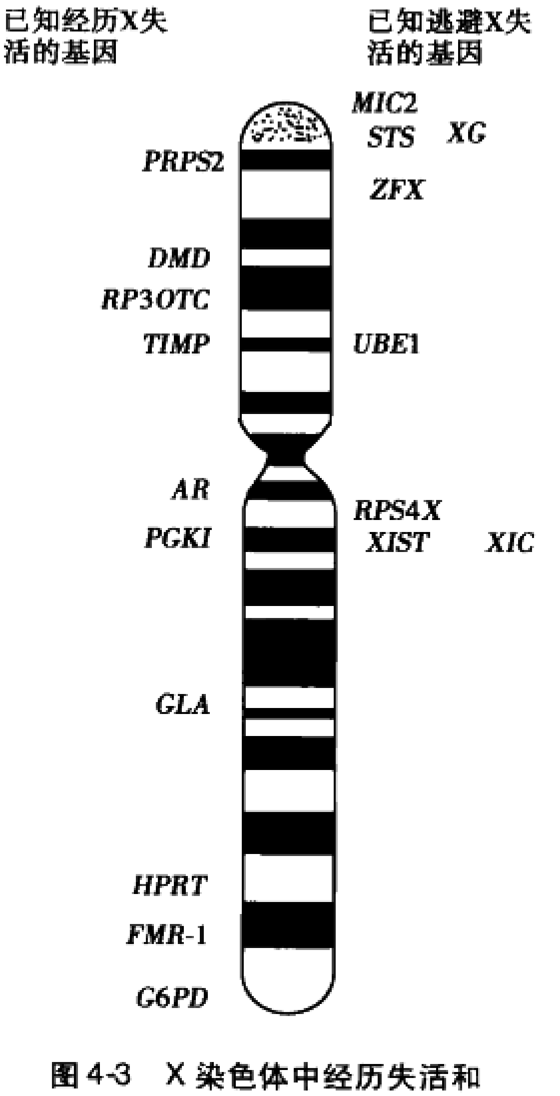


图 4-3 X 染色体中经历失活和逃避失活的基因

女性体细胞的 X 染色体失活通常是随机的，但是也有例外，如 X 染色体发生不平衡型结构异常（包括缺失、重复和等臂染色体）时，失活的总是这条结构异常的 X 染色体，原因可能是一种理性选择，从而避免由于不平衡细胞所引起的临床后果。

非随机 X 失活也可见于 X 与常染色体易位。若易位是平衡型的，则正常的 X 染色体优先失活，2 个染色体的易位部分仍保持活性，这又是一种特殊的选择性，以避免细胞的常染色体基因失活。但是，在平衡携带者的不平衡型子代中，只出现有 X 失活中心的易位产物，故该 X 染色体是永远失活的。这种非随机性失活可缓解而非绝对消除特异染色体缺陷引起的临床后果。在不平衡的 X 与常染色体易位中，正常的 X 染色体有活性，异常的 X 失活。

在女性的 X 与常染色体平衡易位携带者中，如易位点在 X 染色体的某个基因上，则会引起此基因的突变，而另一个正常的 X 染色体处失活状态，不表达，导致细胞中该基因唯一正常拷贝的失活。女性会出现通常只在男性中观察到的 X 连锁表型。

2. Y染色质　正常男性的间期细胞用荧光染料染色后，在细胞核内可出现一强荧光小体，直径为0.3μm左右，称为Y染色质（图4-4）。这是由于Y染色体长臂远端2/3区段为结构异染色质，可被荧光染料染色。这是男性细胞中特有的，女性细胞中不存在。细胞中Y染色质的数目与Y染色体的数目相等。如核型为47，XYY的个体，其间期细胞核中有两个Y染色质。

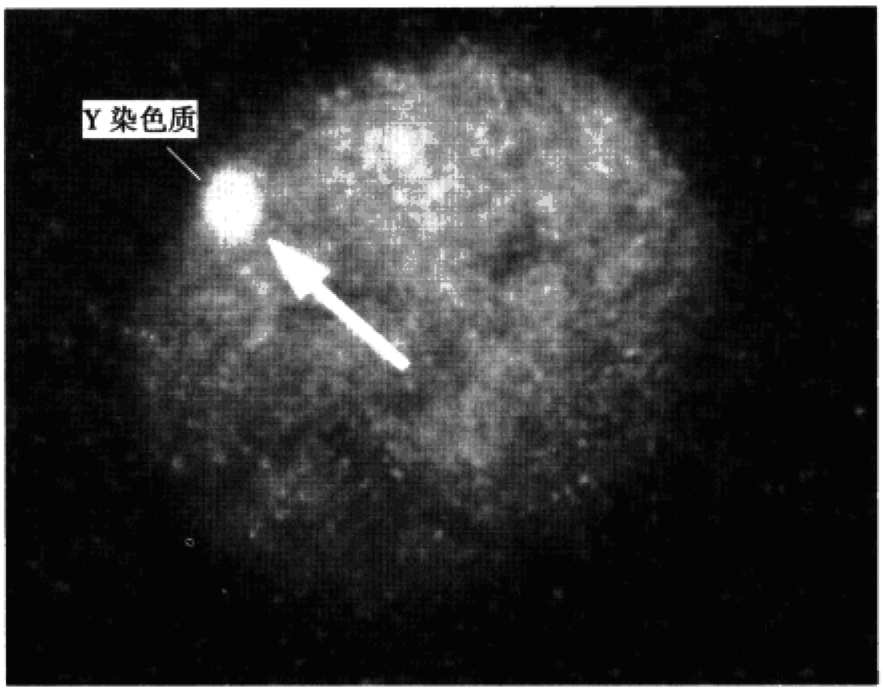


图4-4　男性间期核的Y染色质

二、人类染色体的数目、结构和形态

1. 人类染色体的数目　生物的不同物种其染色体数目各不相同，而同一物种的染色体数目是相对恒定的。例如，果蝇的染色体数目为8，小鼠染色体数为40。染色体数目的恒定对维持物种的稳定性具有重要意义。染色体数目也是物种鉴定的重要标志之一。

在真核生物中，一个正常生殖细胞（配子）中所含的全套染色体称为一个染色体组，其上所包含的全部基因称为一个基因组（genome）。具有一个染色体组的细胞称为单倍体（haploid），以n表示；具有两个染色体组的细胞称为二倍体（diploid），以2n表示。人类正常体细胞染色体数目是46，即2n=46条。正常配子（精子或卵子）中染色体数为23条，即n=23条。

2. 人类染色体的形态、结构　在细胞增殖周期中的不同时期，染色体的形态结构不断的变化着。在有丝分裂中期的染色体的形态是最典型的，可以在光学显微镜下观察，常用于染色体研究和临床上染色体病的诊断。

每一中期染色体都具有两条染色单体（chromatid），互称为姐妹染色单体，它们各含有一条DNA双螺旋链。两条单体之间由着丝粒（centromere）相连接，着丝粒处凹陷缩窄，称初级缢痕（primary constriction）。着丝粒是动粒（kinetochore）形成的位点，并与纺锤体的微管相连，在细胞分裂中与染色体的运动密切相关，失去着丝粒的染色体片段通常不能在分裂后期向两极移动而丢失。着丝粒还含有"卫星"DNA序列，它是一种短串联重复DNA序列，通常具有染色体特异性。着丝粒将染色体划分为短臂（p）和长臂（q）两部分。在短臂和长臂的末端分别有一特化部位称为端粒（telomere）。端粒是一种特殊的蛋白质-DNA结构，含有TTAGGG六核苷酸重复的延伸序列，起着维持染色体形态结构的稳定性和完整性的作用。它可以保护染色体末端不被降解，并防止与其他染色体间的末端融合。端粒长度的缩短与体细胞的老化有关。在某些染色体的长、短臂上还可见凹陷缩窄的部分，称为次级缢痕（secondary constriction）。人类近端着丝粒染色体的短臂末端有一球状结构，称为随体（satellite）。随体柄部为缩窄的次级缢痕。次级缢痕与核仁的形成有关，称为核仁形成区或核仁组织者区（nucleolus organizing region，NOR）。核仁组织者区含有核糖体RNA基因18s和28s的rDNA，其主要功能是转录rRNA，参与核糖体大亚基前体的合成。

染色体上的着丝粒位置是恒定不变的，根据染色体着丝粒的位置可将染色体分为4种类型：①中着丝粒染色体（metacentric chromosome），着丝粒位于或靠近染色体中央。若将染色体全长分为8等份，则着丝粒位于染色体纵轴的1/2～5/8之间，着丝粒将染色体分为长短相近的两个臂；②亚中着丝粒染色体（submetacentric chromosome），着丝粒位于染色体纵轴的5/8～7/8之间，着丝粒将染色体分为长短不同的两个臂；③近端着丝粒染色体（acrocentric chromosome），着丝粒靠近一端，位于染色体纵轴的7/8～末端之间，短臂很短；④端着丝粒染色体（telocentric chromosome），着丝粒位于染色体的

末端，没有短臂。人类染色体只有前三种类型，即中着丝粒染色体、亚中着丝粒染色体和近端着丝粒染色体(图4-5)。

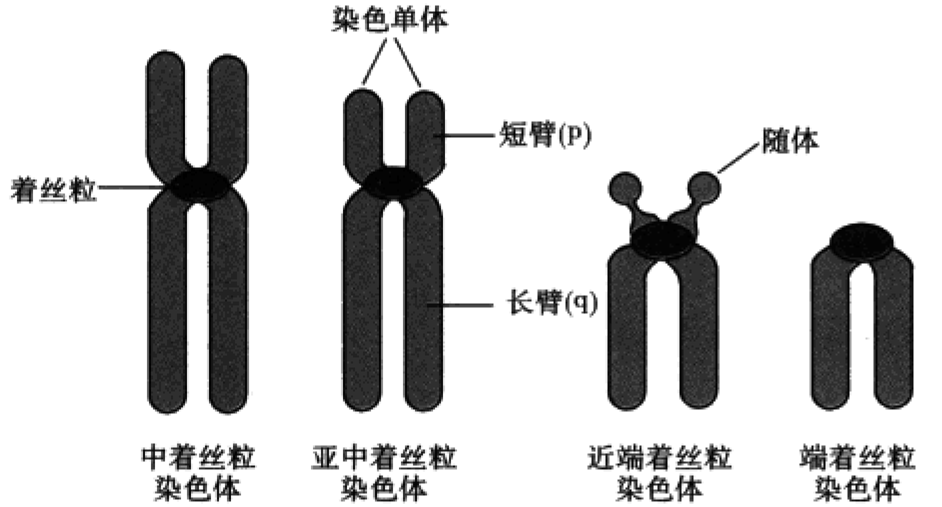


图4-5 染色体的四种类型

三、性染色体与性别决定

人类性别是由细胞中的性染色体所决定的。在人类的体细胞中有23对染色体，其中22对染色体与性别无直接关系，称为常染色体(autosome)。常染色体中的每对同源染色体的形态、结构和大小都基本相同；而另外一对与性别的决定有明显而直接关系的染色体，X染色体和Y染色体，称为性染色体(sex chromosome)。两个性染色体的形态，结构和大小都有明显的差别。X染色体的长度介于C组第6号和第7号染色体之间，而Y染色体的大小与G组第21号和22号染色体相当。男性的性染色体组成为XY，而在女性细胞中的性染色体组成为XX，即男性为异型性染色体，女性为同型性染色体。这种性别决定方式为XY型性别决定。因此，在配子发生时，男性可以产生两种精子，含有X染色体的X型精子和含有Y染色体的Y型精子，两种精子的数目相等；而女性则由于细胞中有两条相同的X染色体，只能形成一种含有X染色体的卵子。受精时，X型精子与卵子结合，形成性染色体组成为XX的受精卵，将来发育成为女性；而Y型精子与卵子结合则形成性染色体组成为XY的受精卵，发育成为男性。所以人类的性别是精子和卵子在受精的瞬间决定的，确切地说是由精子决定的。在自然状态下，不同的精子与卵子的结合是随机的，因此人类的男女比例大致保持1:1。

很显然，性别决定实际上是由精子中所带有的性染色体是X染色体还是Y染色体决定的，而X染色体和Y染色体在人类性别决定中的作用并不相等。一个个体的体细胞中无论有几条X染色体，只要有Y染色体就决定男性表型(睾丸女性化患者除外)。因为Y染色体的短臂上有一个决定男性性别的基因——睾丸决定因子(testis-determining factor，TDF)基因，*TDF*基因是性别决定的关键基因。性染色体异常的个体，如核型为47，XXY或48，XXXY等，他们的表型是男性，但却是一个不正常的男性。没有Y染色体的个体，其性腺发育基本上是女性特征，即使只有一条X染色体如核型为45，X的个体，其表型也是女性，但却是一个表型异常的女性。

1990年Sinclair等发现了一个新基因，被命名为性别决定区域Y，即*SRY*(sex-determining region Y)，并且认为是TDF的最佳候选基因。*SRY*基因位于Y染色体短臂末端，其产物为SRY蛋白，决定睾丸的形成。而且有越来越多的证据表明*SRY*和人类性别决定确有关系。

近年的研究表明，*SRY*可能作为睾丸发育的启动者，它是性别决定中最重要的基因，但不是唯一决定基因。*SRY*必须激活睾丸分化途径上其他的基因才能使内生殖器进行正常发育。所以可能有多个基因影响性别的决定。性别决定与分化是一个相当复杂的过程，可能涉及了性染色体和常染色体上多个基因的协同作用。

四、染色体的研究方法

对人类染色体的研究已有很长的历史,1888 年德国解剖学者 Waldeyer 根据细胞有丝分裂和生殖细胞减数分裂观察到的现象,提出了染色体这一名称。但是由于人类染色体数目较多,并且由于当时的技术和方法的限制,对染色体的研究受到一定的影响。尤其是染色体数目的研究结果很不一致。1923 年 Painter 提出了染色体数目为 2n = 48 的观点,这个结论一直被多数学者所承认。直到 1956 年,华裔学者蒋有兴(Joe Hin Tjio)和 Albert Levan 应用纺锤丝抑制剂——秋水仙碱和低渗技术,在流产的胎儿肺组织培养中发现这些细胞的染色体数目是 46 条,而不是 48 条。英国学者 Charles Ford 和 John Hamerton 的研究结果支持了他们的结论。从此肯定了人类染色体数目为 2n = 46。这标志着现代细胞遗传学的开始。

(一) 染色体核型分析

一个体细胞中的全部染色体,按其大小、形态特征顺序排列所构成的图像就称为核型(karyotype)。将待测细胞的核型进行染色体数目、形态特征的分析,确定其是否与正常核型完全一致,称为核型分析(karyotype analysis)。

1. 人类染色体非显带核型　非显带染色体核型是指按常规染色方法所得到的染色体标本,用

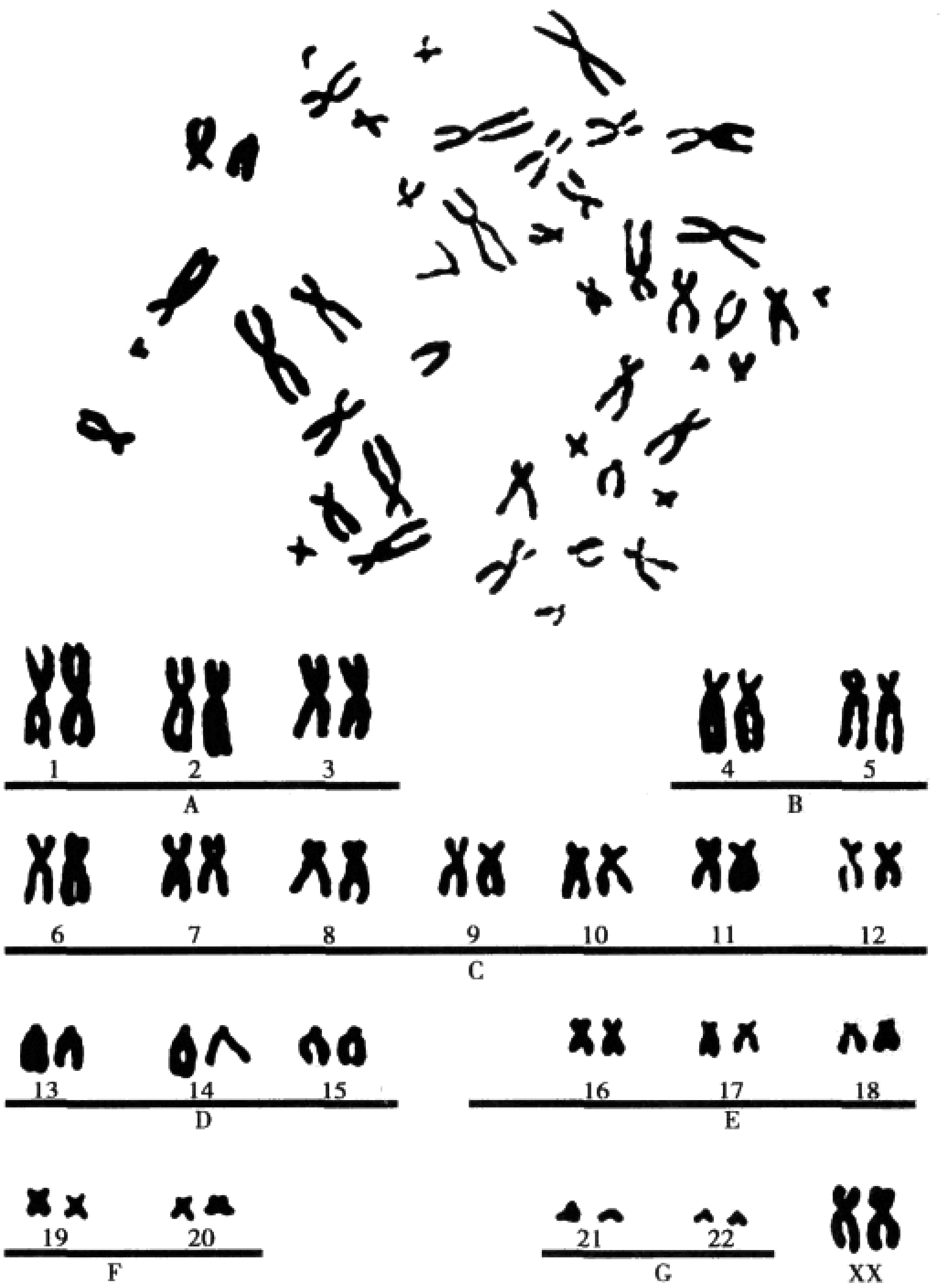


图 4-6　人类染色体非显带核型(46,XX)

Giemsa 染色，染色体除着丝粒和次级缢痕外，整条染色体均匀着色，因此，很难准确鉴别多数组内染色体的序号（图 4-6）。

1960 年在美国丹佛、1963 年在英国伦敦、1966 年在美国芝加哥召开过三次国际会议，确定和制定了人类有丝分裂染色体的识别、编号、分组以及核型描述（包括染色体数目和结构异常的核型描述）等一套统一的标准命名系统。主要根据染色体长度和着丝粒的位置等，将人的体细胞 46 条染色体进行配对、顺序排列、编号。1 ~ 22 号为常染色体（euchromosome），是男女共有的 22 对染色体。其余一对随男女性别而异，为性染色体（sex chromosome），女性为 XX，男性为 XY。将这 23 对染色体分为 A、B、C、D、E、F、G 7 个组，A 组最大，G 组最小。X 染色体列入 C 组，Y 染色体列入 G 组（见表 4-1）。

表 4-1　人类核型分组与各组染色体形态特征（非显带标本）

组号	染色体号	大小	着丝粒位置	次级缢痕	随体	可鉴别程度
A	1 ~ 3	最大	中（1、3 号）	1 号常见	无	可鉴别
			亚中（2 号）			
B	4 ~ 5	次大	亚中		无	难鉴别
C	6 ~ 12；X	中等	亚中	9 号常见	无	难鉴别
D	13 ~ 15	中等	近端		有	难鉴别
E	17 ~ 18	小	中（16 号）	16 号常见	无	16 号可鉴别
			亚中（17、18 号）		无	17、18 难鉴别
F	19 ~ 20	次小	中		无	难鉴别
G	21 ~ 22；Y	最小	近端		21、22 号有	难鉴别
					Y 无	

核型的描述包括两部分内容：①染色体总数，②性染色体的组成，两者之间用“，”分隔开。正常女性核型描述为：46，XX。正常男性核型描述为：46，XY。在正常核型中，染色体是成对存在的，每对染色体在形态结构、大小和着丝粒位置上基本相同，其中一条来自父方的精子，一条来自母方的卵子，称为同源染色体（homologous chromosome）。而不同对染色体彼此称为非同源染色体。

2. 人类染色体显带核型　非显带染色体标本是用 Giemsa 染色液使染色体着色，不能将每一条染色体本身的特征完全显示出来，因此，只能根据各染色体的大致特征（大小、着丝粒位置）来识别染色体，即使是最有经验的细胞遗传学家，也只能较准确地识别出 1、2、3、16 号和 Y 等几条染色体，对 B、C、D、F 和 G 组的染色体，只能识别出属于哪一组，而对组内相邻号的染色体之间很难区分，并且，对于染色体所发生的一些结构畸变，例如易位、倒位和微小的缺失等均不能检出，这对染色体异常，特别是结构畸变的研究与临床应用受到极大的限制。因此，从 1959 年 Lejeune 发现第一例人类染色体病至 1968 年的 10 年中，人们只发现了 10 多种染色体异常综合征，并且主要是染色体数目异常的病例。

1968 年瑞典细胞化学家 Caspersson 等应用荧光染料氮芥喹吖因（quinacrine mustard，QM）处理染色体后，在荧光显微镜下可观察到染色体沿其长轴显示出一条条宽窄和亮度不同的横纹，即染色体的带（band）。这一显带技术称 Q 显带（Q banding），所显示的带纹称为 Q 带（Q band）（图 4-7a）。显带技术可将人类的 24 种染色体显示出各自特异的带纹，称为带型（banding pattern）。随后又出现了其他几种染色体显带技术：①G 显带（G banding）：是将染色体标本用碱、胰蛋白酶或其他盐溶液处理后，再用吉姆萨（Giemsa）染色，染色体上出现与 Q 带相类似的带纹，在普通显微镜下，可见染色深浅相间的带纹，称 G 带（G band）（图 4-7b）。G 带与 Q 带相对应，即在 Q 显带的亮带的相应部位，被 Giemsa 染成深带，而在 Q 显带中暗带的相应部位被染成浅带。G 显带方法简便，带纹清晰，染色体标本可以长期保存，因此被广泛用于染色体病的诊断和研究。②R 显带（R banding）：用盐溶液处理标本

后，再用 Giemsa 染色，显示出与 G 带相反的带，即 G 显带中的深带在 R 显带中为浅带，G 显带中的浅带在 R 显带中为深带，称反带（reverse band）或 R 带（R band）（图 4-7c）。③T 显带（T banding）：将染色体标本加热处理后，再用 Giemsa 染色可使染色体末端区段特异性深染，称 T 带（T band）。④C 显带（C banding）：用 NaOH 或 $Ba(OH)_2$ 处理标本后，再用 Giemsa 染色，可使着丝粒和次级缢痕的结构异染色质部分深染，如 1、9、16 号染色体的次级缢痕以及 Y 染色体长臂远端的 2/3 的区段，所显示的带纹称 C 带（C band）（图 4-7d）。C 显带可用于检测 Y 染色体、着丝粒区以及次级缢痕区的变化。⑤N显带：用硝酸银染色，可使染色体的随体及核仁形成区（NOR）呈现出特异性的黑色银染物，这种银染色阳性的 NOR 称为 Ag-NOR。据研究表明，Ag-NOR 的可染性取决于它的功能活性，即具转录活

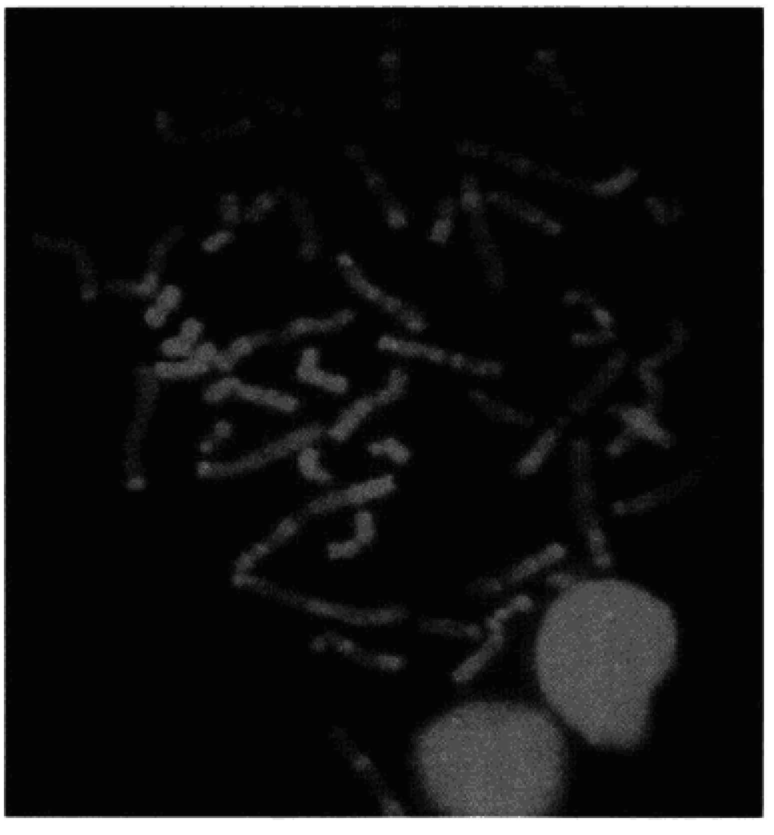


图 4-7a　Q 显带



图 4-7b　G 显带

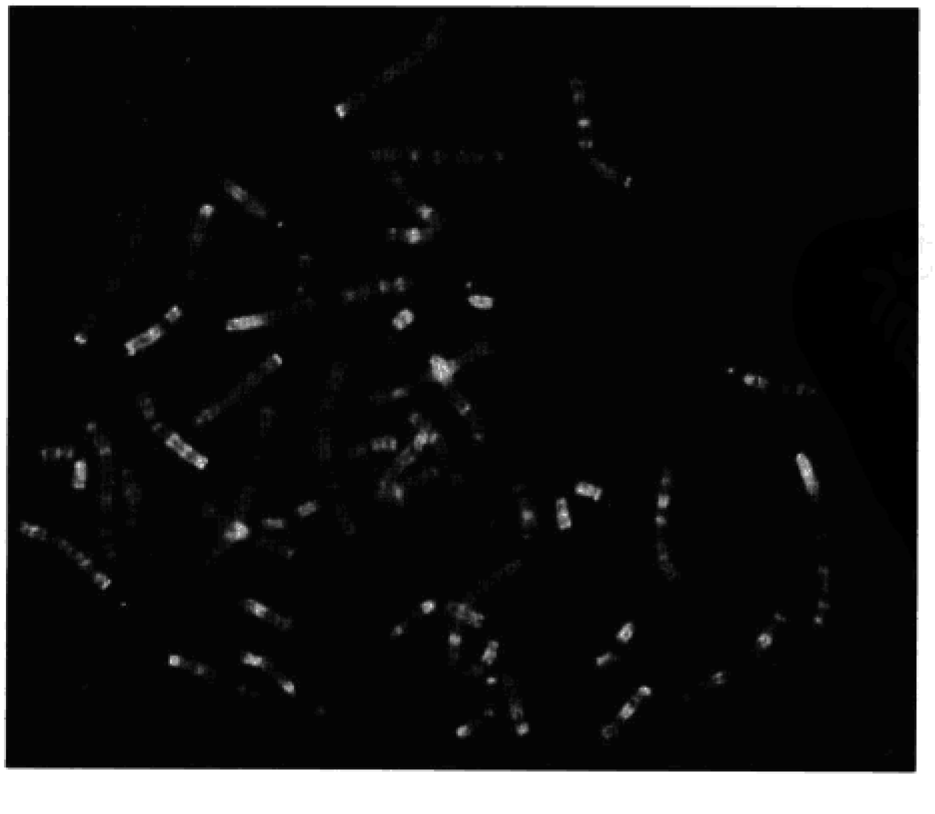


图 4-7c　R 显带染色体核型

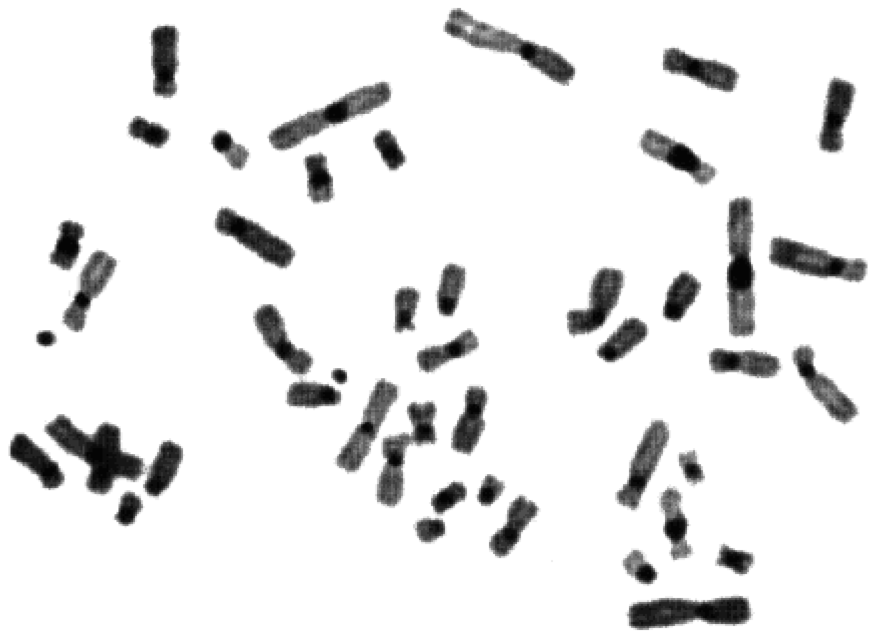


图 4-7d C 显带染色体(女性)

性的NOR着色，但受染物质不是次级缢痕本身，而是附近与 rDNA 转录有关的一种酸性蛋白。

用 Q 显带、G 显带和 R 显带等染色体显带方法，可使染色体沿其长轴显示出明暗或深浅相间的带纹，而每一条染色体都有其独特而恒定的带纹，这就构成了每条染色体的带型。同源染色体的带型基本相同，不同对的染色体的带型各不相同。这为识别每条染色体提供了分析基础，通过染色体显带核型分析，我们可以准确的识别每一条染色体以及其所发生的各种变异。

3. G 显带染色体的识别 目前，G 显带核型分析已成为临床常规应用的染色体病诊断的手段之一。下面介绍正常人类体细胞中期染色体 G 带的带型的识别要点。

在进行 G 带带型描述时，“深带”表示被 Giemsa 着色的带纹，“浅带”表示不着色或基本不着色的带纹。“浓”、“淡”表示深带着色的强度。用“近侧段”、“中段”、“远侧段”来表示距离着丝粒的远近(图 4-8，图 4-9)。

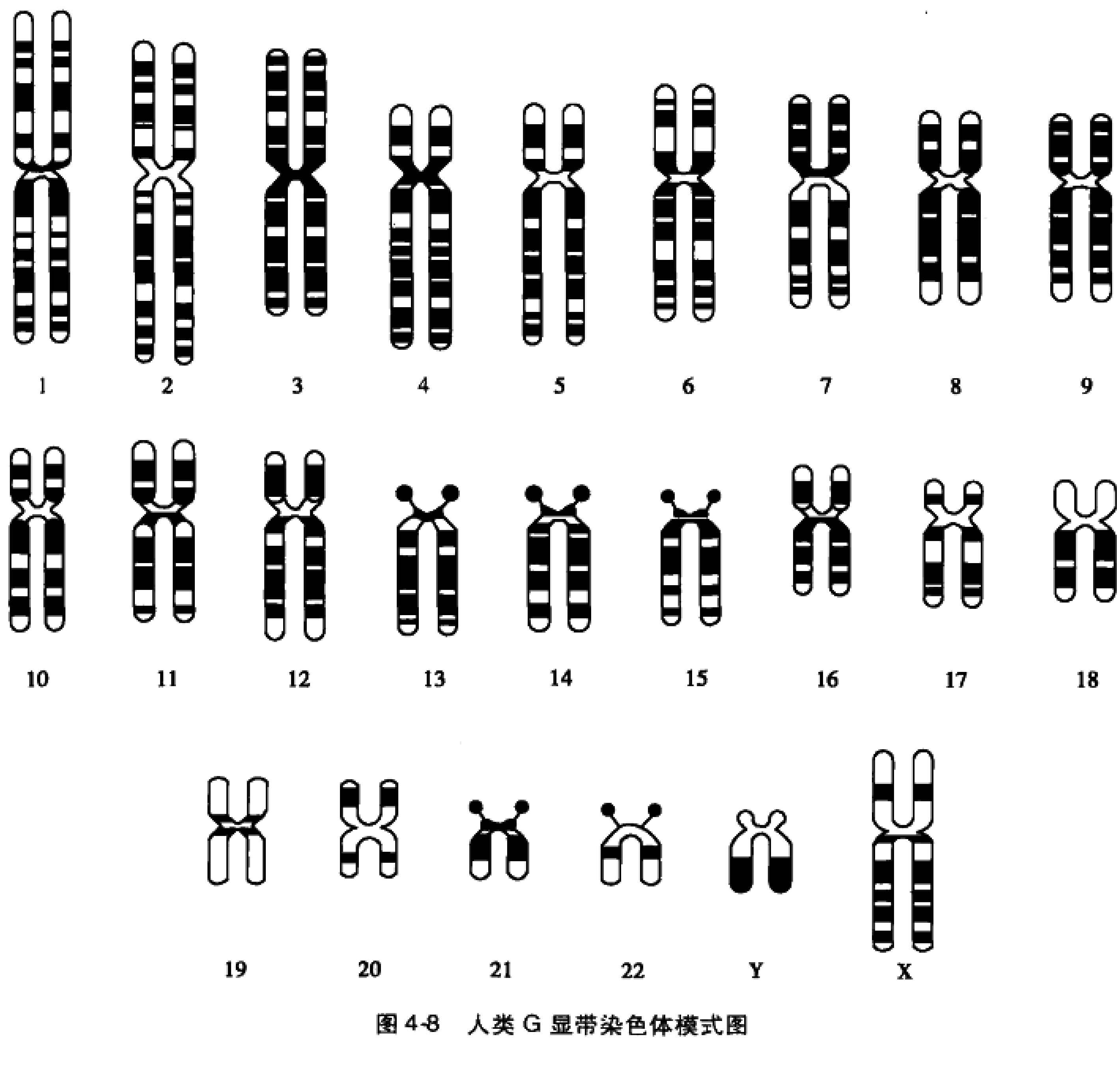


图 4-8 人类 G 显带染色体模式图

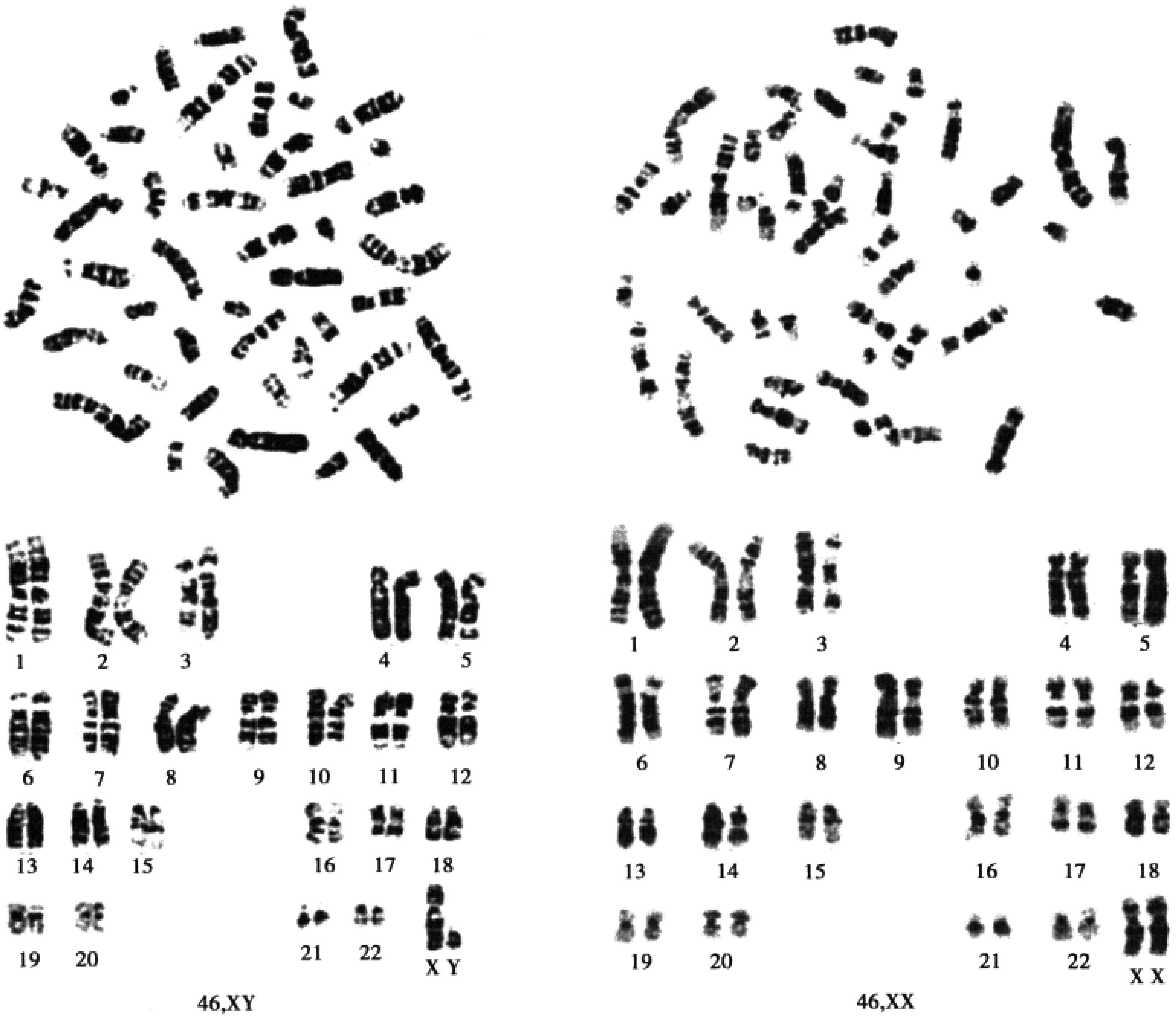


图 4-9　人类染色体 G 显带核型

A 组染色体：包括 1～3 号染色体，是长度最长的染色体，其中 1、3 号为中着丝粒染色体，2 号染色体的着丝粒为亚中近中部。

B 组染色体：包括 4～5 号染色体，为最大的亚中着丝粒染色体。

C 组染色体：包括 6～12 号和 X 染色体，为中等大小的亚中着丝粒染色体，X 染色体的大小类似于 7 号染色体。

D 组染色体：包括 13～15 号染色体，为中等大小的带有随体的近端着丝粒染色体。

E 组染色体：包括 16～18 号染色体，为较小的中着丝粒（16 号）和亚中着丝粒（17、18 号）染色体。

F 组染色体：包括 19～20 号染色体，为小的中着丝粒染色体。

G 组染色体：包括 21～22 号和 Y 染色体。其中 21、22 号为小的带随体的近端着丝粒染色体，Y 染色体无随体。

4. 人类细胞遗传学国际命名体制　随着细胞遗传学研究方法的广泛使用，遗传学家们在丹佛（1960 年）、伦敦（1963 年）、芝加哥（1966 年）和巴黎（1970 年）召开了国际人类细胞遗传学会议，统一了细胞遗传学的命名原则。国际人类细胞遗传学命名委员会于 1978 年第一次出版了《人类细胞遗传学国际命名体制》（An International System for Human Cytogenetics Nomenclature, ISCN），规定了正常及异常核型的命名格式和原则。这样对显带染色体有了一个统一的识别和描述的标准，有利于相互交流。此后，ISCN 的专家委员先后在 1981、1985、1990 和 1995 年召开了会议，对人类染色体命名规

则进一步修改并出版了新的版本。其中1981年版是人类高分辨显带的命名体制，1991年版是肿瘤细胞遗传学的命名体制，1995年版首次刊登了分子细胞遗传学命名原则和格式。

每条显带染色体根据ISCN规定的界标(landmark)划分为若干个区，每个区(region)又包括若干条带(band)。界标是确认每一染色体上具有重要意义的、稳定的、有显著形态学特征的指标，包括染色体两臂的末端、着丝粒和某些显著的带。两相邻界标之间为区。每一条染色体都是由一系列连贯的带组成，没有非带区。它借助其亮-暗或深-浅的着色强度，清楚地与相邻的带相区别。

每一染色体都以着丝粒为界标，分成短臂(p)和长臂(q)。区和带的序号均从着丝粒为起点，沿着每一染色体臂分别向长臂、短臂的末端依次编号为1区、2区、……，以及1带、2带、……。界标所在的带属于此界标以远的区，并作为该区的第1带。被着丝粒一分为二的带，分别归属于长臂和短臂，分别标记为长臂的1区1带和短臂的1区1带(图4-10)。

描述一特定带时需要写明以下4个内容：①染色体序号；②臂的符号；③区的序号；④带的序号。例如：1p31表示第1号染色体短臂3区1带。

应用染色体显带技术可以识别染色体细微的结构异常。为了能够简明的描述这些异常的核型，1977年在斯德哥尔摩，1981年在巴黎召开的国际会议上议定的《人类细胞遗传学命名的国际体制》(ISCN，1978，1981)，制定了统一的命名符号和术语(表4-2)。

（二）高分辨显带染色体

人类中期染色体的带纹数较少。一套单倍体染色体带纹数仅有320条带。70年代后期，由于技术的改进，可以从早中期、前中期、晚前期细胞得到更长、带纹更丰富的染色体。一套单倍体染色体即可显示550～850条或更多的带纹。即在中期染色体原有的带纹上分出更多更细的带称为亚带。这种染色体称为高分辨显带染色体(high resolution banding chromosome，HRBC)(图4-11)。

“人类细胞遗传学高分辨显带命名的国际体制(1981)(ISCN 1981)”的模式图，显示了大约具有550～850条带的高分辨带。高分辨显带的命名方法是在原带之后加“.”，并在“.”之后写亚带的序

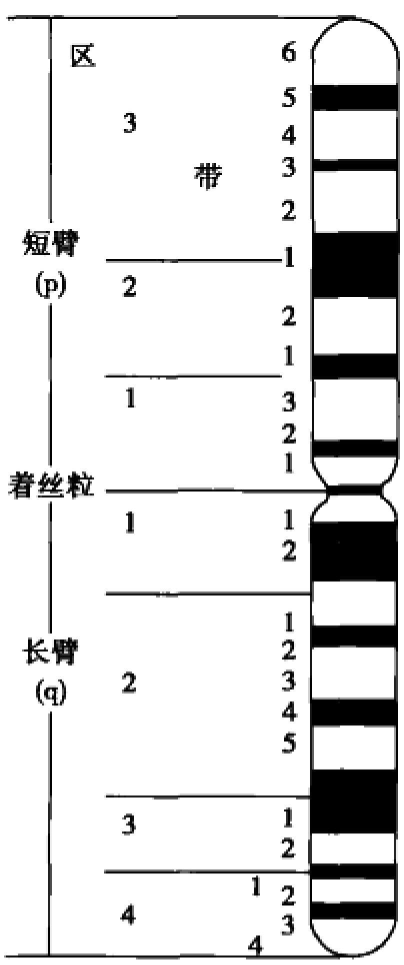


图4-10 人类显带染色体界标、区、带的示意图

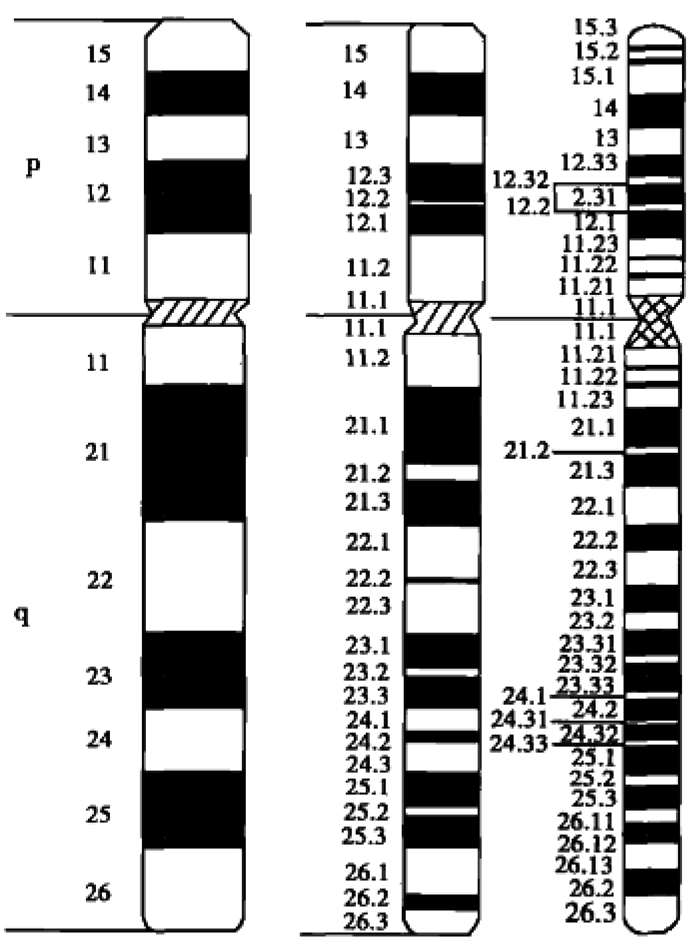


图4-11 人类10号染色体高分辨显带模式图

表4-2　核型分析中常用的符号和术语

符号术语	意　义	符号术语	意　义
A～G	染色体组的名称	+或-	在染色体和组号前面，表示染色体或组内染色体增加或减少；在臂或结构后面，表示这个臂或结构的增加或减少
1～22	常染色体序号		
→	从…到…		
/	表示嵌合体染色体		
ace	无着丝粒断片（见f）		
?	分类或情况不明	mat	母源的
cen	着丝粒	min	微小体
chi	异源嵌合体	mn	众数
:	断裂	mos	嵌合体
::	断裂与重接	p	短臂
ct	染色单体	pat	父源的
del	缺失	ph	费城染色体
der	衍生染色体	pro	近侧
dic	双着丝粒	psu	假
dir	正位	q	长臂
dis	远侧	qr	四射体
dmin	双微体	r	环状染色体
dup	重复	rcp	相互易位
t	易位	rea	重排
end	（核）内复制	rac	重组染色体
f	断片	rob	罗伯逊易位
fem	女性	s	随体
fra	脆性部位	tan	串联易位
g	裂隙	ter	末端
h	次级缢痕	tr	三射体
i	等臂染色体	tri	三着丝粒
ins	插入	var	可变区
inv	倒位	mar	标记染色体
mal	男性		

号。例如：原来的1p36带被分为三个亚带，分别命名为1p31.1，1p31.2，1p31.3，即表示1号染色体短臂3区1带第1亚带、第2亚带、第3亚带。若亚带再分成更细的带，称为次亚带，如1p31.3再分成三条次亚带，则写为1p31.31，1p31.32，1p31.33。

染色体高分辨显带能为染色体及其所发生的畸变提供更多细节，有助于发现更多、更细微的染色体结构异常，使染色体发生畸变的断裂点定位更加准确。因此，这一技术无论在临床细胞遗传学、分子细胞遗传学的检查上，或者是在肿瘤染色体的研究和基因定位上都有广泛的应用价值。

（三）姐妹染色单体交换（sister chromatid exchange，SCE）：指一条染色体的两条姐妹染色单体在细胞内可自发地或在某些因素作用下在同一位置同时发生断裂，并互换片段后重新接合的现象。因为是在同一位置上发生的对等同源片段的交换，染色体的形态和带型都没有发生改变。由于SCE在普通染色标本上无法观察到，因此，需用特殊染色方法来检测。常用的方法是在细胞培养液中加入

5-溴脱氧尿嘧啶核苷(5-Bromodeoxy uridine，BrdU)，进行细胞培养后制片，经特殊方法处理后 Giemsa 染色，然后在显微镜下选择处于第二个分裂周期的中期分裂相进行观察，可见一条染色体的两条姐妹染色单体出现明显的差别染色，一条深染，一条浅染，并可见到姐妹染色单体互换现象(图 4-12)。

产生 SCE 的机理尚未完全阐明，但它显然与 DNA 的损伤和修复过程有关，因此检测 SCE 的发生率对各种理化因素导致遗传物质损伤的研究，肿瘤病因的研究以及对细胞周期的研究等方面都有重要的意义。



图 4-12　姐妹染色单体交换

(四) 分子细胞遗传学技术

分子细胞遗传学(molecular cytogenetics)是传统细胞遗传学与分子遗传学相结合的一门学科。分子细胞遗传学的发展始于荧光原位杂交(FISH)，从最早的单色 FISH 发展成为现在的 24 色 FISH 的过程中，所采用的方法包括多色 FISH(multiplex FISH，M-FISH)、光谱核型分析(spectral karyotyping，SKY)、物种交叉色带(cross species color banding，RxFISH)、CCK(color changing karyotyping)和比较基因组杂交(comparative genomic hybridization，CGH)等。这些新兴技术将 FISH 的敏感性和特异性与传统的细胞遗传染色体分析技术结合起来，能筛查整个基因组的染色体异常，特别是在检测复杂的肿瘤染色体畸变方面显示了独特的优势。分子细胞遗传学在医学遗传学的研究中起到了推动性的作用，并在临床诊断中发挥了日益重要的作用。

1. 荧光原位杂交技术　荧光原位杂交(fluorescence in situ hybridization，FISH)技术是在原位杂交技术的基础上发展起来的。其基本原理与原位杂交一样，是利用 DNA 碱基互补配对的特点，在体外的一定条件下，使同源的 DNA 链或 DNA-RNA 单链结合成双链。FISH 使用荧光标记的 DNA、RNA 或与 mRNA 互补的 cDNA 探针和染色体或基因杂交，从而在中期染色体、间期核、组织切片、裂殖细胞或配子细胞上检测 DNA 顺序。

FISH 方法是先将靶 DNA 及其周围物质固定于玻片上，通过加热和甲酰胺处理使靶 DNA 双链变性成单链，同时使双链 DNA 探针变性形成单链，然后在适当条件下使单链探针 DNA 与单链的靶 DNA 结合或杂交形成新的双链 DNA。如果探针 DNA 与待测 DNA 上的同源序列(靶序列)互补结合，即可在该染色体上原位(靶序列的位置)显示杂交信号。这种探针不仅能与中期染色体进行杂交，而且还能与非分裂细胞固定的间期核直接进行杂交。标准细胞遗传学技术需要分裂细胞和阻断在分裂中期的染色体，而 FISH 技术将染色体分析扩展到间期细胞。因而 FISH 技术具有快速、安全、经济、灵敏度高和特异性强等优点，而且标本可以长期保存而不失活。因此，FISH 技术已广泛应用于医学生物学研究及临床诊断中，如基因定位、基因扩增及染色体畸变的检测等。

2. DNA 纤维荧光原位杂交　DNA 纤维荧光原位杂交(DNA-fiber-FISH)是一种可目视的高分辨基因组制图技术。其杂交及检测分析步骤和染色体的 FISH 基本相同，但染色体的 FISH 分辨率为1～2Mb，所以只能将染色体定位到亚带水平，不适于定位克隆研究。而 DNA Fiber FISH 的分辨率可达到 10kb，探针长度在 1～300kb 范围内，在定位克隆中具有重要作用。因此，DNA 纤维荧光原位杂交已成功应用于人类基因组制图、染色质结构分析以及染色体病、肿瘤等遗传性疾病的分析研究中。

这一技术的基本原理是利用碱溶液或甲醛溶液处理待测细胞，使间期核染色质的组织结构松散，

将染色质(丝)从核骨架中释放出来，在载玻片上制备出DNA纤维，然后将不同颜色荧光物质标记的特异DNA探针，分别杂交到靶细胞的DNA纤维上，根据杂交的结果判断各个探针的定位、方向以及各探针之间的物理距离和重叠程度，确定DNA微小缺失与重复，作出基因定位。

3. 染色体涂染技术　染色体涂染(chromosome painting probes)是将荧光原位杂交和染色体原位杂交相结合而建立起来的。即用单链DNA或Cot-1封闭基因组DNA重复序列，以减少非特异性杂交信号，增强特异性杂交的信号强度，并用染色体特异性DNA库作为探针池，用不同的荧光涂染整条染色体或染色体特异区段，从而使待测的整条染色体或染色体特异区段显示出发荧光的杂交信号，根据结果作出分析诊断。

用于染色体涂染的探针可以是整条染色体，也可以是染色体的特异区段。探针的来源主要有：①含有人单条染色体的人-啮齿类(human-rodent)体细胞杂种组织融合的产物；②荧光激活的流式细胞仪(fluorescence-activated cell sorter，FACS)分离整条染色体DNA，并以载体克隆、PCR扩增制备探针；③经显微切割得到染色体或染色体片段，再经PCR扩增制作的探针池。染色体涂染方法有正向(forward)涂染和逆向(reverse)涂染两种。正向涂染是将正常的探针杂交到待检测的异常标本上。逆向涂染则是将分离到的异常染色体制备成探针，然后杂交到正常标本上。

目前染色体涂染技术已成功应用于染色体数目和结构异常分析、不同物种间的同源性比较和白血病及其他肿瘤的染色体诊断和研究。

4. 比较基因组杂交　比较基因组杂交技术是在FISH的基础上建立起来的。探针是整个基因组DNA，而不是一个点或一个区域。可在全部染色体或在染色体亚带水平上，对不同基因组之间DNA序列拷贝进行检测和定位。主要用于确定未知区域的DNA扩增或缺失，如染色体不平衡片段的识别和染色体重排的研究、相关种属间或同一种属内不同个体之间基因组差异的研究等。尤其是CGH回避了实体肿瘤染色体制备的难题，这是其他FISH方法难以替代的。目前，CGH已广泛应用于肿瘤遗传的研究中。

5. 光谱核型分析　与常规的FISH不同，光谱核型分析(SKY)除了杂交外，还对染色体进行不同颜色组合的标记，使每条染色体分别显示出不同的颜色谱。如使用两种不同颜色标记探针，可以观察到三种颜色的靶染色体(A，B或A+B，即2^2-1)。若使用三种不同颜色标记探针，则可以观察到7种不同颜色的靶染色体(2^3-1)，以此类推，使用5种不同颜色的荧光标记探针，就可以产生$2^5-1=31$种颜色的靶染色体，比代表24条染色体的24种颜色还多。通常使用如Cy3、Cy5、Cy5.5、Spectrum Green和TexasRed等5种不同颜色的荧光染料所产生的不同颜色组合对每一条染色体进行标记，然后通过频谱遥感连接CCD(charge-coupled-device)摄像仪和电脑等装置，使每对同源染色体都显示出其独特的染色谱，从而可以一目了然地观察全部染色体的变化。光谱核型分析(SKY)方法可以同时分析24条染色体，因而可以一次性地将复杂畸变或不同的染色体畸变检测出来。

由于每对同源染色体都有独特的颜色，因此可应用SKY技术，观察所有染色体的结构改变，识别标记染色体以及常规染色体核型分析不能识别的复杂或隐蔽的染色体重排。研究表明，SKY可以准确地检测大于1.5Mb的染色体异常。目前，SKY已经应用于产前诊断和肿瘤的细胞遗传学诊断中。

(五) 人类染色体的多态性

在正常健康人群中，存在着各种染色体的恒定的微小变异，包括结构、带纹宽窄和着色强度等。这类微小而恒定的变异是按照孟德尔方式遗传的，通常没有明显的表型效应或病理学意义，称为染色体多态性(chromosomal polymorphism)。

染色体多态性常见于以下部位：①Y染色体的长度变异，这种变异存在着种族差异。主要变异部位是Y染色体长臂结构异染色质区，即长臂远端约2/3区段的长度变异。如果Y染色体大于F组或大于第18号染色体，称为"长Y"、"大Y"或"巨Y"，描述为Yq^+；若Y染色体的长度为G组染色体长度的1/2以下，称"小Y"染色体，描述为Yq^-，但这种现象比较罕见。②D组、G组近端着丝粒染色体的短臂、随体及随体柄部的次级缢痕区(NOR)的变异。表现为随体的有无、大小及重复(双随体)等；

短臂、次级缢痕的增长或缩短。③第1、9和16号染色体次级缢痕的变异,表现为次级缢痕的有无或长短的差异。此外,在1、9和16号染色体的着丝粒异染色质区也可出现多态性的倒位。

染色体的多态性变异主要发生在结构异染色质区,因此一般没有明显的表型效应和病理学意义,也就是说一般没有不良的临床后果。但现在有研究报道,某些多态现象与临床症状有关。这说明染色体多态性与表型效应之间的关系问题,还有待于进一步的研究探讨。

染色体多态现象是一种较稳定的结构变异,可以在显微镜下观察检查,并且它是按孟德尔方式遗传的,它以一定的遗传方式传给下一代,因此可以作为一种遗传标记,应用于临床和研究工作。

染色体多态性的应用:

(1) 染色体的多态性可用于追溯染色体(包括额外染色体或异常染色体)的来源。比如在产前诊断中,进行羊水细胞或绒毛细胞检查时可根据染色体多态现象来鉴定胎、母细胞,判断有无母体细胞的污染;可以根据21号染色体的短臂、随体、次级缢痕以及显带着色强度等多态性特征来追溯额外的21号染色体来自父方或母方,确定21三体综合征患者的额外染色体的来源。

(2) 法医中可用以进行亲权鉴定。通过检查子女和父母(或可能的父母)的染色体,根据染色体多态性标记的异同,可以帮助判断子女与其父母的真实关系,进行亲权鉴定。

(3) 染色体多态性变异可作为一种标记,进行不同种族或民族人群中的遗传学研究。

第二节 染色体畸变

染色体畸变(chromosomal aberration)是指体细胞或生殖细胞内染色体发生异常的改变。畸变的类型和可能引起的后果在细胞不同周期和个体发育不同阶段不尽相同。染色体畸变可分为数目畸变和结构畸变两大类。其中染色体的数目畸变又可分为整倍性改变和非整倍性改变两种。结构畸变主要有缺失、重复、插入、易位和倒位等。当一个个体细胞有两种或两种以上的不同核型的细胞系时,这个个体就被称为嵌合体。无论数目畸变,还是结构畸变,其实质是涉及染色体上基因群的增减或位置的转移,使遗传物质发生了改变,可以导致染色体异常综合征或染色体病。

一、染色体畸变发生的原因

导致染色体畸变的因素有多种,归纳起来可以分为以下几种:化学因素、物理因素、生物因素和母亲年龄。

(一) 化学因素

许多化学物质,如一些化学药品、农药、毒物和抗代谢药等,都可引起染色体畸变。据调查,某些化工厂的工人由于长期接触苯、甲苯等,出现染色体数目异常和发生染色体断裂的频率远高于一般人群。农药中的除草剂和杀虫的砷制剂等都是染色体畸变的诱变剂。

1. 药物　某些药物特别是一些抗肿瘤药物、保胎及预防妊娠反应的药物,均可引起人类染色体畸变或产生畸胎。已有研究证实,环磷酰胺、氮芥、白消安(马利兰)、氨甲蝶呤、阿糖胞苷等抗癌药物均可导致染色体畸变。抗痉挛药物苯妥英钠可引起人淋巴细胞多倍体细胞数增高。

2. 农药　许多化学合成的农药可以引起人类细胞染色体畸变。某些有机磷农药也可使染色体畸变率增高,如美曲磷酯(敌百虫)类农药。

3. 工业毒物　工业毒物如苯、甲苯、铝、砷、二硫化碳、氯丁二烯、氯乙烯单体等,都可以导致染色体畸变。长期接触这些有害毒物的工人,其染色体的畸变率增高。

4. 食品添加剂　某些食品的防腐剂和色素等添加剂中所含的化学物质也可以引起人类染色体发生畸变。如硝基呋喃基糖酰胺AF-2、环己基糖精等。

（二）物理因素

在自然空间存在的各种各样的射线都可对人体产生一定的影响，但其剂量极微，故影响不大。但大量的电离辐射对人类具有极大的潜在危险。例如放射线物质爆炸后散落的放射性尘埃、医疗上所用的放射线等，对人体都有一定的损害。工业放射性物质的污染也可引起细胞染色体的改变。细胞受到电离辐射后，可引起细胞内染色体发生异常。畸变率随射线剂量的增高而增高。最常见的畸变类型有断裂、缺失、双着丝粒染色体、易位、核内复制、不分离等，这些畸变都可使个体的性状出现异常。射线的作用包括对体细胞和生殖细胞两方面，如果一次照射大剂量的射线，可在短期内引起造血障碍而死亡。长期接受射线治疗或从事放射工业的人员，由于微小剂量的射线不断积累，会引起体细胞或生殖细胞染色体畸变。有实验证明，受照射卵细胞中染色体不分离的频率明显高于未受照射组。同时还发现，这一现象在年龄较大的小鼠中更为明显。还有人报道，受到过电离辐射的母亲生育先天愚型（一种染色体异常所导致的疾病）患儿的风险明显增高。

（三）生物因素

导致染色体畸变的生物因素包括两个方面：一是由生物体产生的生物类毒素所致，二是某些生物体如病毒本身可引起染色体畸变。真菌毒素具有一定的致癌作用，同时也可引起细胞内染色体畸变。病毒也可引起宿主细胞染色体畸变，尤其是那些致癌病毒。其原因主要是由于影响 DNA 代谢。当人体感染某些病毒，如风疹病毒、乙肝病毒、麻疹病毒和巨细胞病毒时，就有可能引发染色体的畸变。如果用病毒感染离体培养细胞将会出现各种类型的染色体异常。

（四）母亲年龄

当母亲年龄增大时，所生子女的体细胞中某一序号染色体有三条的情况要多于一般人群。母亲年龄越大（大于 35 岁），生育先天愚型患儿的危险性就越高。这与生殖细胞老化及合子早期所处的宫内环境有关。一般认为，生殖细胞在母体内停留的时间越长，受到各种因素影响的机会越多，在以后的减数分裂过程中，容易产生染色体不分离而导致染色体数目异常。

二、染色体数目畸变

（一）整倍体改变

如果染色体的数目变化是单倍体（n）的整倍数，即以 n 为基数，整倍地增加或减少，则称为整倍体（euploid）。超过二倍体的整倍体称为多倍体（polyploid）。

在 2n 的基础上，如果增加一个染色体组，也就是增加一个 n，则为 3n，即三倍体（triploid）。若在 2n 的基础上增加两个 n，则为 4n，即四倍体（tetraploid）。以此类推。三倍体以上的又统称为多倍体。如果在 2n 的基础上减少一个染色体组，则称为单倍体（haploid）。

在人类中已知有三倍体和四倍体的个体，但只有极少数三倍体的个体能存活到出生，存活者多为 2n/3n 的嵌合体。有调查资料表明，在自发流产的胎儿中，有染色体畸变的约占 50%，其中，三倍体占 18%，四倍体占 5%。可见在流产的胎儿中三倍体是常见的类型。一般认为，三倍体胎儿易于流产的原因是在胚胎发育过程的细胞有丝分裂中，形成三极纺锤体，因而造成染色体在细胞分裂中期、后期时的分布和分配紊乱，最终导致子细胞中染色体数目异常，从而严重干扰了胚胎的正常发育而导致流产。四倍体比三倍体更为罕见，往往是四倍体和二倍体（4n/2n）的嵌合体，或在流产的胚胎中发现。

整倍体改变的机制主要有：双雌受精、双雄受精、核内复制和核内有丝分裂等。

1. 双雌受精（digyny） 一个二倍体的异常卵子与一个正常的精子发生受精，从而产生一个三倍体的合子。在卵细胞发生的第二次减数分裂过程中，次级卵母细胞由于某种原因未形成第二极体，因此应分给第二极体的染色体组仍留在卵细胞中，使该卵细胞成为异常卵细胞。当它与一个正常的精子结合后，就将形成含有三个染色体组的合子，即三倍体。可形成 69，XXX 或 69，XXY 两种核型的受精卵（图 4-13a）。

2. 双雄受精(diandry) 一个正常的卵子同时与两个正常的精子发生受精。由于每个精子都带有一个染色体组，所以当两个精子同时进入一个卵细胞时，就将两个染色体组同时带入了这一卵细胞，所形成的合子内则含有三个染色体组，即三倍体。可形成69,XXX、69,XXY和69,XYY三种类型的受精卵(图4-13b)。

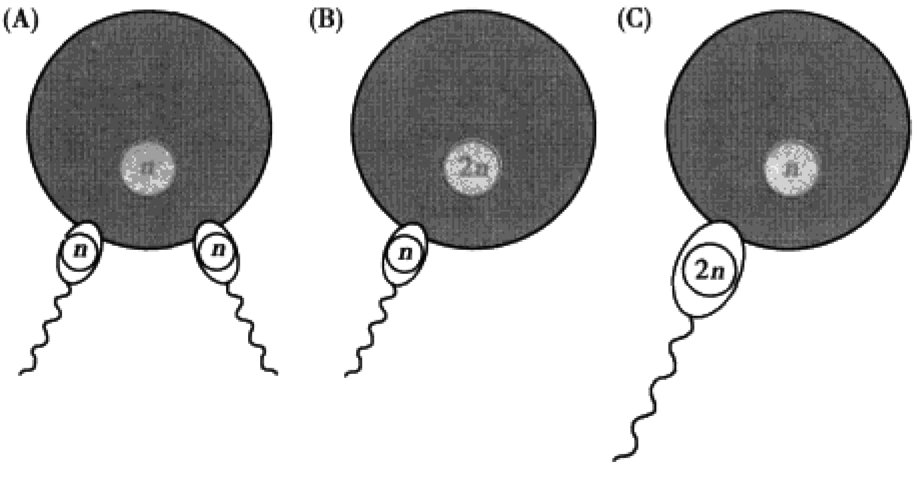


图4-13a 三倍体形成
(A)、(C)双雄受精；(B)双雌受精



图4-13b 四倍体的形成

3. 核内复制(endoreduplication) 指在一次细胞分裂时，DNA不是复制一次，而是复制了两次，而细胞只分裂了一次。这样形成的两个子细胞都是四倍体。这是肿瘤细胞常见的染色体异常特征之一。

4. 核内有丝分裂(endomitosis) 在细胞分裂时，染色体正常复制了一次，但至分裂中期时，核膜仍未破裂、消失，也无纺锤体的形成，因此，细胞分裂未能进入后期和末期，没有细胞质的分裂，结果细胞内含有四个染色体组，形成了四倍体。

归纳来说，三倍体的形成原因可为双雌受精或双雄受精；四倍体形成的主要原因是核内复制和核内有丝分裂。

(二) 非整倍体改变

一个体细胞的染色体数目增加或减少了一条或数条，称非整倍体(aneuploid)。这是临床上最常见的染色体畸变类型。发生非整倍体改变后，会产生亚二倍体(hypodiploid)、超二倍体(hyperdiploid)等。亚二倍体即在2n的基础上，减少了一条或几条染色体，可写作2n－m(注：m<n)；超二倍体即在2n的基础上，增加了一条或几条染色体，可写作2n＋m(注：m<n)。

1. 亚二倍体 当体细胞中染色体数目少了一条或数条时，称为亚二倍体(hypodiploid)。若某对染色体少了一条(2n－1)，细胞染色体数目为45，即构成单体型(monosomy)。临床上常见的有21号、22号和X染色体的单体型，核型为45,XX(XY),－21、45,XX(XY),－22和45,X。核型为45,X的个体往往是由于X染色体的丢失所致，具有这种核型的个体，多在胚胎期流产，只有少数存活的个体。由于缺少一条X染色体，具有性腺发育不全等临床症状。

2. 超二倍体 当体细胞中染色体数目多了一条或数条时，称为超二倍体(hyperdiploid)。在超二倍体的细胞中某一同源染色体的数目不是2条，而是3条、4条……

若某对染色体多了一条(2n＋1)，细胞内染色体数目为47，即构成三体型(trisomy)。这是人类染色体数目畸变中最常见、种类最多的一类畸变。例如，在常染色体病中，除了第17号染色体尚未有三体型的病例报道外，其余的染色体三体型均有报道，但是由于染色体的增加，特别是较大染色体的增加，将造成基因组的严重失衡而破坏或干扰胚胎的正常发育，故绝大部分常染色体三体型核型只见于早期流产的胚胎。少数三体型病例可以存活至出生，但多数寿命不长，并伴有各种严重畸形。

三体型以上的统称为多体型(polysomy)。多体型常见于性染色体中，如性染色体四体型(48,

XXXX;48,XXXY;48,XXYY)和五体型(49,XXXXX;49,XXXYY)等。如果患者细胞中一对同源染色体同时缺失,即减少了一对同源染色体(2n－2),称为缺体型(nullosomy)。人类缺体型尚未见报道,这种核型的个体是不能存活的。

3. 假二倍体　有时细胞中某染色体数目发生了异常,其中有的增加,有的减少,而增加和减少的染色体数目相等,结果染色体总数不变,还是二倍体数(46 条),但不是正常的二倍体核型,则称为假二倍体(pseudodiploid)

4. 嵌合体　一个个体内同时存在两种或两种以上核型的细胞系,这种个体称嵌合体(mosaic)。如46,XX/47,XXY;45,X/46,XX 等。嵌合体可以是数目异常之间、结构异常之间以及数目和结构异常之间的嵌合。

(三) 非整倍体的产生机制

非整倍体的产生原因,多数是在性细胞成熟过程或受精卵早期卵裂中,发生了染色体不分离或染色体丢失。

1. 染色体不分离(nondisjunction)　在细胞分裂进入中、后期时,如果某一对同源染色体或姐妹染色单体彼此没有分离,而是同时进入一个子细胞,结果所形成的两个子细胞中,一个将因染色体数目增多而成为超二倍体,另一个则因染色体数目减少而成为亚二倍体,这个过程称为染色体不分离。染色体不分离可以发生在细胞的有丝分裂过程,也可以发生在配子形成时的减数分裂过程。

(1) 染色体不分离发生在受精卵的卵裂早期的有丝分裂过程中。卵裂早期某一染色体的姐妹染色单体不分离,可导致产生由两种细胞系或三种细胞系组成的嵌合体。不分离发生在第一次卵裂,则形成具有两个细胞系的嵌合体,一个为超二倍体细胞系,一个为亚二倍体细胞系。不分离发生在第二次卵裂以后,即形成具有三个或三个以上细胞系的嵌合体(45/46/47)(图 4-14)。不分离发生得越晚,正常二倍体细胞系的比例越大,临床症状也相对较轻。

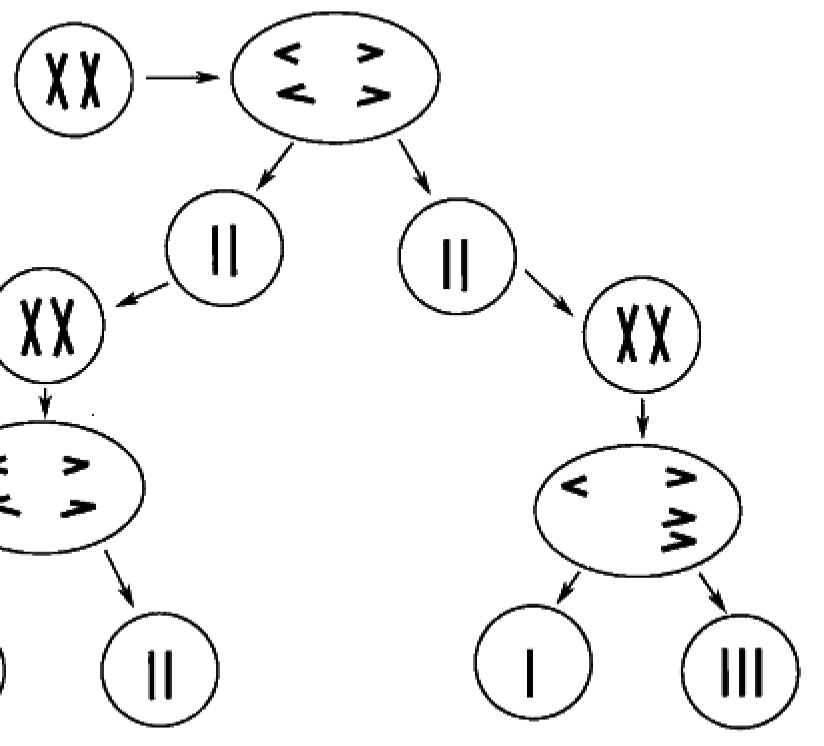


图 4-14　有丝分裂时姐妹染色体不分离

(2) 减数分裂时发生染色体不分离。染色体不分离发生在减数分裂Ⅰ,使得某一对同源染色体不分离,同时进入一个子细胞核,所形成的配子中,一半将有 24 条染色体(n＋1),另一半将有 22 条(n－1)。与正常配子受精后,将形成超二倍体或亚二倍体。若在减数分裂Ⅱ发生染色体不分离,所形成的配子的染色体数将有以下几种情况:1/2 为 n;1/4 为(n＋1),1/4 为(n－1)。它们与正常配子受精后,得到相应的二倍体、超二倍体、亚二倍体(图 4-15)。

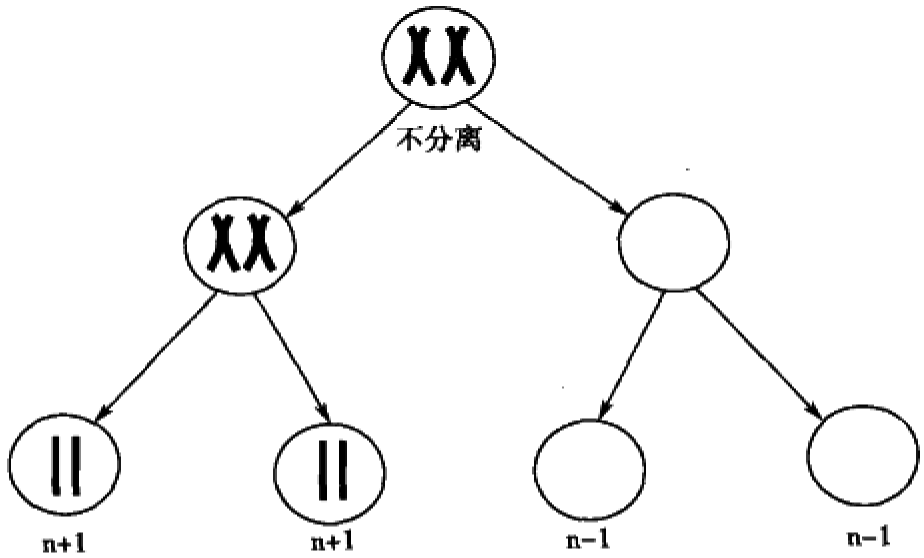


图 4-15a 减数分裂Ⅰ染色体不分离

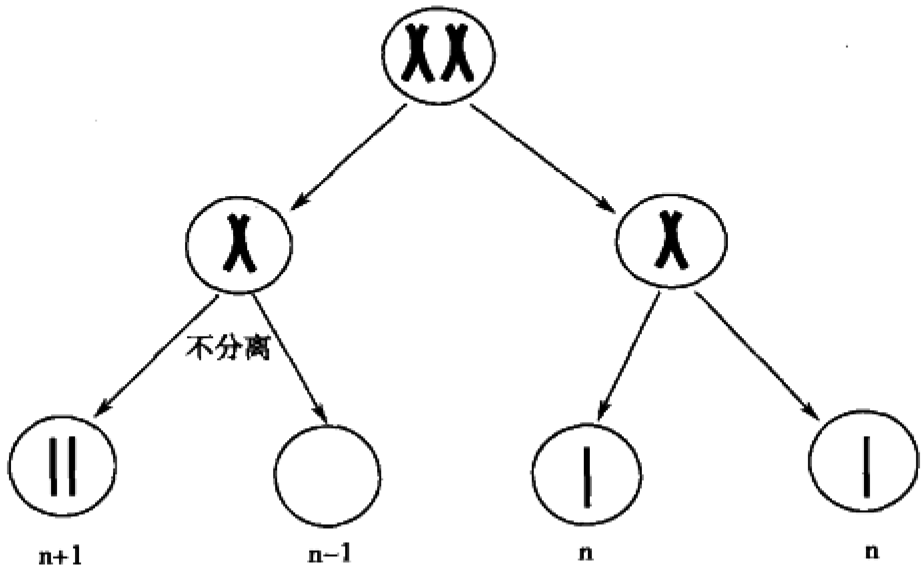


图 4-15b 减数分裂Ⅱ染色体不分离

2. 染色体丢失 染色体丢失(chromosome lose)又称染色体分裂后期延滞(anaphase lag),在细胞有丝分裂过程中,某一染色体未与纺锤丝相连,不能移向两极参与新细胞的形成;或者在移向两极时行动迟缓,滞留在细胞质中,造成该条染色体的丢失而形成亚二倍体。染色体丢失也是嵌合体形成的一种方式。

按照 ISCN(1978),非整倍体的描述方法为"染色体总数,性染色体组成,+(-)畸变染色体序号"。例如某一核型中的 18 号染色体多了一条,可描述为:47,XX(XY),+18;少了一条 22 号染色体则描述为 45,XX(XY),-22;若是少了一条 X 染色体,可描述为 45,X 或 45,XO。

三、染色体结构畸变

染色体结构畸变的发生受多种因素的影响,如物理因素、化学因素、生物因素和遗传因素等。在这些因素的作用下,首先是染色体发生断裂(breakage),然后是断裂片段的重接(rejoin)。断裂的片段如果在原来的位置上重新接合,称为愈合或重合(reunion),即染色体恢复正常,不引起遗传效应。如果染色体断裂后未在原位重接,也就是断裂片段移动位置与其他片段相接或者丢失,则可引起染色体结构畸变,又称染色体重排(chromosomal rearrangement)。

(一) 染色体结构畸变的描述方法

人类细胞遗传学命名的国际体制(ISCN)制定了有关人类染色体以及染色体畸变等的命名方法。结构畸变染色体核型的描述方法有简式和详式两种。在简式中,对染色体结构的改变只用其断点来表示。按国际命名规定(表 4-3),应依次写明染色体总数,性染色体组成,然后用一个字母(如 t)或三

联字符号(如 del)写明重排染色体的类型,其后的第一个括弧内写明染色体的序号,第二个括弧写明区号、带号以表示断点。在详式中,除了简式中应写明的内容外,与简式有所不同,即在最后一个括弧中不是只描述染色体的断裂点,而是描述重排染色体带的组成。

表 4-3　显带染色体结构畸变符号

畸变类型	畸变符号及其意义
环形	46,XY,r(2)(p21q31) 46,XY,r(2)(p21→q31) 第 2 号染色体短臂 2 区 1 带与长臂 3 区 1 带断裂点相接成环形
等臂	46,X,i(Xq) 46,X,i(X)(qter→cen→qter) 一条正常 X 染色体和一条 X 长臂等臂染色体,后者是从 X 长臂末端到着丝粒再到长臂末端止
末端缺失	46,XX,del(1)(q21) 46,XX,del(1)(pter→q21:) 第 1 号染色体长臂 2 区 1 带处断裂,其远端部分缺失,保留短臂末端到长臂 2 区 1 带处止
中间缺失	46,XX,del(1)(q21q31) 46,XX,del(1)(pter→q21::q31→qter) 第 1 号染色体长臂 2 区 1 带处和 3 区 1 带处断裂,中间部分缺失,它们又再重新相接,所以保留从短臂末端到长臂 2 区 1 带,再与 3 区 1 带相接到长臂末端止
相互易位	46,XY,t(2;5)(q21;q31) 46,XY,t(2;5)(2pter→2q21::5q31→5qter;5pter→5q31::2q21→2qter) 第 2 号染色体长臂 2 区 1 带断裂,其远端部易位到第 5 号染色体,而第 5 号染色体长臂 3 区 1 带处断裂,其远端部分易位到 2 号上,重组形成 2 条新的染色体。一条是自 2 号短臂末端到 2 号长臂 2 区 1 带处,再接于 5 号染色体的长臂 3 区 1 带到 5 号长臂末端;另一条是自 5 号短臂末端到 5 号长臂 3 区 1 带处,再接于 2 号染色体的长臂 2 区 1 带到 2 号长臂末端
臂内倒位	46,XY,inv(2)(p13p24) 46,XY,inv(pter→p24::p13→p24::p13→qter) 断裂和连接发生于 2 号染色体短臂 1 区 3 带和 2 区 4 带处,这部分片段倒位后重接,使 2 区 4 带处和 1 区 3 带连接,而 1 区 3 带处则与 2 区 4 带处连接,造成这部分顺序颠倒,但其着丝粒类型未变
臂间倒位	46,XY,inv(2)(p21q31) 46,XY,inv(2)(pter→p21::q31→p21::q31→qter) 断裂和连接发生在 2 号染色体短臂 2 区 1 带和长臂 3 区 1 带处,位于这些带间断片顺序颠倒后重接,使长臂 3 区 1 带连接于短臂的 2 区 1 带处,而另一侧短臂的 2 区 1 带则与长臂的 3 区 1 带处相接,造成 2p21-2q31 间的顺序颠倒了。由于该断片涉及着丝粒,因此重接后的染色体着丝粒类型可以有所变化

(二) 染色体结构畸变的类型及其产生机制

临床上常见的染色体结构畸变有:缺失、重复、易位、倒位、环状染色体和等臂染色体等。

1. 缺失(deletion)　染色体片段的丢失,使位于这个片段的基因也随之发生丢失。按染色体断点的数量和位置可分为末端缺失和中间缺失两类:①末端缺失(terminal deletion)是指染色体的臂发生断裂后,未发生重接,无着丝粒的片段不能与纺锤丝相连而丢失。如图 4-16a 所示,第 1 号染色体的长臂的 2 区 1 带发生断裂,其远侧段(q21→qter)丢失。这条染色体是由短臂的末端至长臂的 2 区 1 带所构成。这种结构畸变的简式描述为:46,XX(XY),del(1)(q21);详式描述为:46,XX(XY),del(1)(pter→q21:);②中间缺失(interstitial deletion)指一条染色体的同一臂上发生了两次断裂,两个断点之间的片段丢失,其余的两个断片重接。如图 4-16b 所示,3 号染色体长臂上的 q21 和 q25 发生断裂和重接,这两断点之间的片段丢失。这种结构畸变的简式描述为:46,XX(XY),del(3)(q21q25);

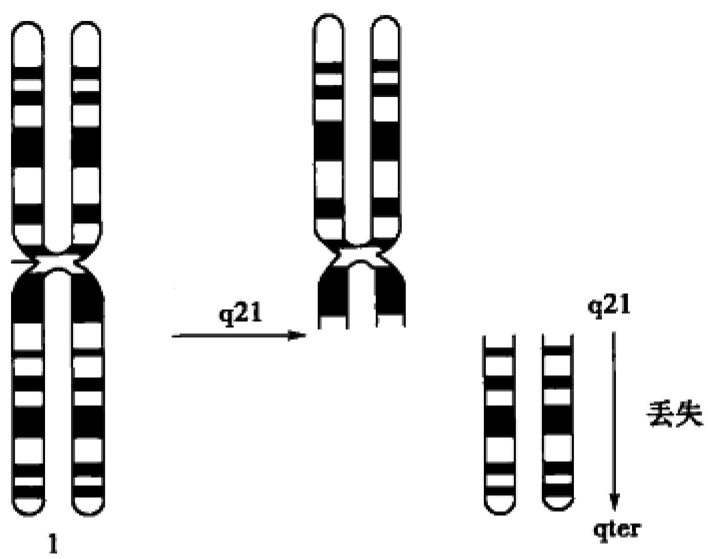


图 4-16a 末端缺失

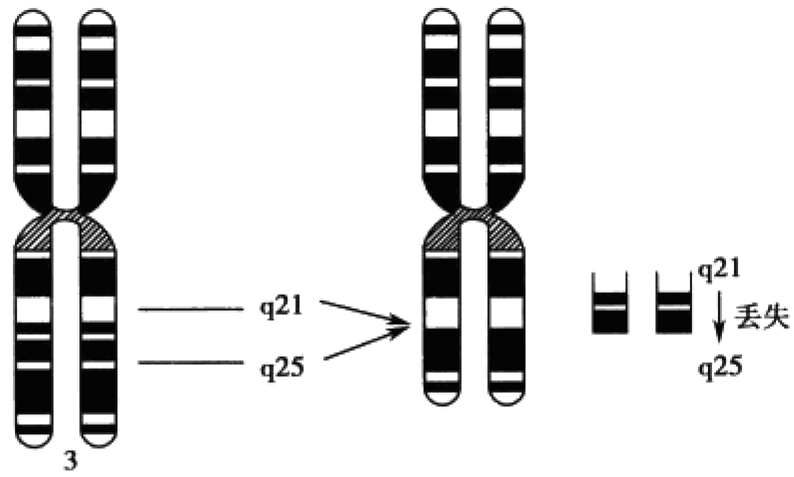


图 4-16b 中间缺失

详式描述为:46,XX(XY),del(3)(pter→q21::q25→qter)。

2. 重复(duplication)一个染色体上某一片段增加了一份或一份以上的现象,使这些片段的基因多了一份或几份。发生的原因是同源染色体之间的不等交换或染色单体之间的不等交换以及同源染色体片段的插入等。

3. 倒位(inversion) 是某一染色体发生两次断裂后,两断点之间的片段旋转 180 度后重接,造成染色体上基因顺序的重排。染色体的倒位可以发生在同一臂(长臂或短臂)内,也可以发生在两臂之间,分别称为臂内倒位和臂间倒位:①臂内倒位(paracentric inversion):一条染色体的某一臂上同时发生了两次断裂,两断点之间的片段旋转 180 度后重接。例如 1 号染色体 p22 和 p34 同时发生了断裂,两断点之间的片段倒转后重接,形成了一条臂内倒位的染色体(图 4-17a)。这种结构畸变的简式描述为:46,XX(XY),inv(1)(p22p34);详式描述为:46,XX(XY),inv(1)(pter→p34::p22→p34::p22→qter);②臂间倒位(pericentric inversion):一条染色体的长、短臂各发生了一次断裂,中间断片颠倒后重接,则形成了一条臂间倒位染色体。如 4 号染色体的 p15 和 q21 同时发生了断裂,两断点之间的片段倒转后重接,形成了一条臂间倒位染色体(图 4-17b)。这种结构畸变的简式描述为:46,XX(XY),inv(4)(p15q21);详式描述为:46,XX(XY),inv(4)(pter→p15::q21→p15::q21→qter)。

具有臂间倒位染色体的个体称为倒位携带者(inversion carrier)。这种个体一般外表正常,但染

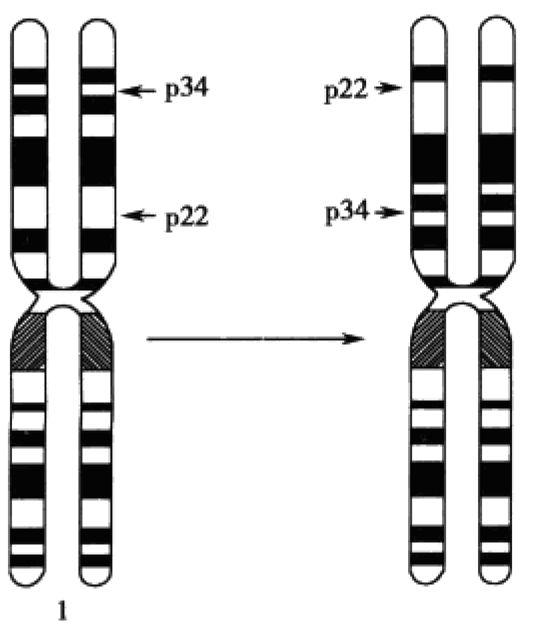


图 4-17a 1 号染色体短臂臂内倒位图解

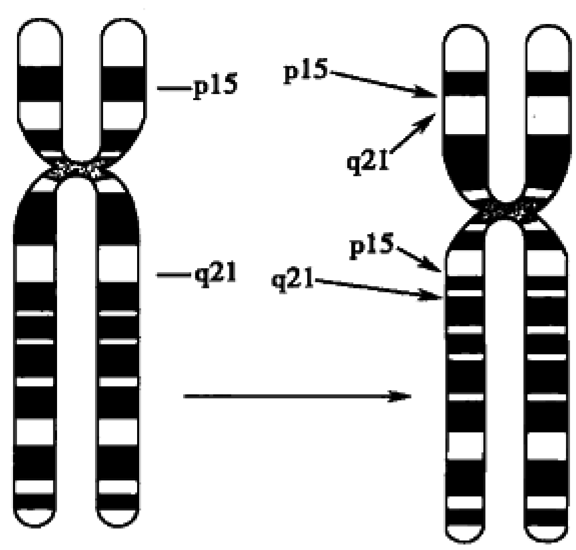


图 4-17b 染色体臂间倒位

色体发生倒位后，其结构发生了重排，形成了重排染色体，这种重排染色体在形成生殖细胞的减数分裂Ⅰ的前期中，同源染色体发生联会配对时形成特有的倒位环。如果在倒位环内发生交换，理论上形成4种不同的配子，一种具有正常染色体，一种具有倒位染色体，其余两种均带有部分重复和部分缺失的染色体。

4. 易位（translocation）　一条染色体的断片移接到另一条非同源染色体的臂上，这种结构畸变称为易位。常见的易位方式有相互易位、罗伯逊易位和插入易位等。①相互易位（reciprocal translocation）两条染色体同时发生断裂，断片交换位置后重接，形成两条衍生染色体（derivation chromosome）。当相互易位仅涉及位置的改变而不造成染色体片段的增减时，则称为平衡易位。如4号染色体长臂3区5带和6号染色体长臂2区1带同时发生了断裂，两断片交换位置后重接，形成两条衍生染色体（图4-18）。这种结构畸变的简式描述为：46，XX（XY）；t（4；6）（q35；q21）；详式描述为：46，XX（XY），t（4；6）（4pter→4q35：：6q21→6qter；6pter→6q21：：4q35→4qter）；②罗伯逊易位（Robertsonian translocation）又称着丝粒融合（centric fusion）。这是发生于近端着丝粒染色体的一种易位形式（图4-19）。当两个近端着丝粒染色体在着丝粒部位或着丝粒附近部位发生断裂后，二者的长臂在着丝粒处接合在一起，形成一条衍生染色体。两个短臂则构成一个小染色体，小染色体往往在第二次分裂时丢失，这可能是由于其缺乏着丝粒或者是由于其完全由异染色质构成所致。由于丢失的小染色体几乎全是异染色质，而由两条长臂构成的染色体上则几乎包含了两条染色体的全部基因。因此，罗伯逊易位携带者虽然只有45条染色体，但表型一般正常，在形成配子的时候会出现异常，造成胚胎死亡而流产或出生先天畸形患儿。如14号染色体短臂的1区1带（14p11）和21号染色体的长臂的1区1带（21q11）同时发生了断裂，两条染色体带有长臂的断片相互连接，即在着丝粒部位融合，形成的衍生染色体包含了21号染色体的21q11→qter节段和14号染色体14p11→qter节段，其余的部分均丢失。③插入易位（insertional translocation），两条非同源染色体同时发生断裂，但只有其中一条染色体的片段插入到另一条染色体的非末端部位。只有发生了三次断裂时，才可能发生插入易位（图4-20）。

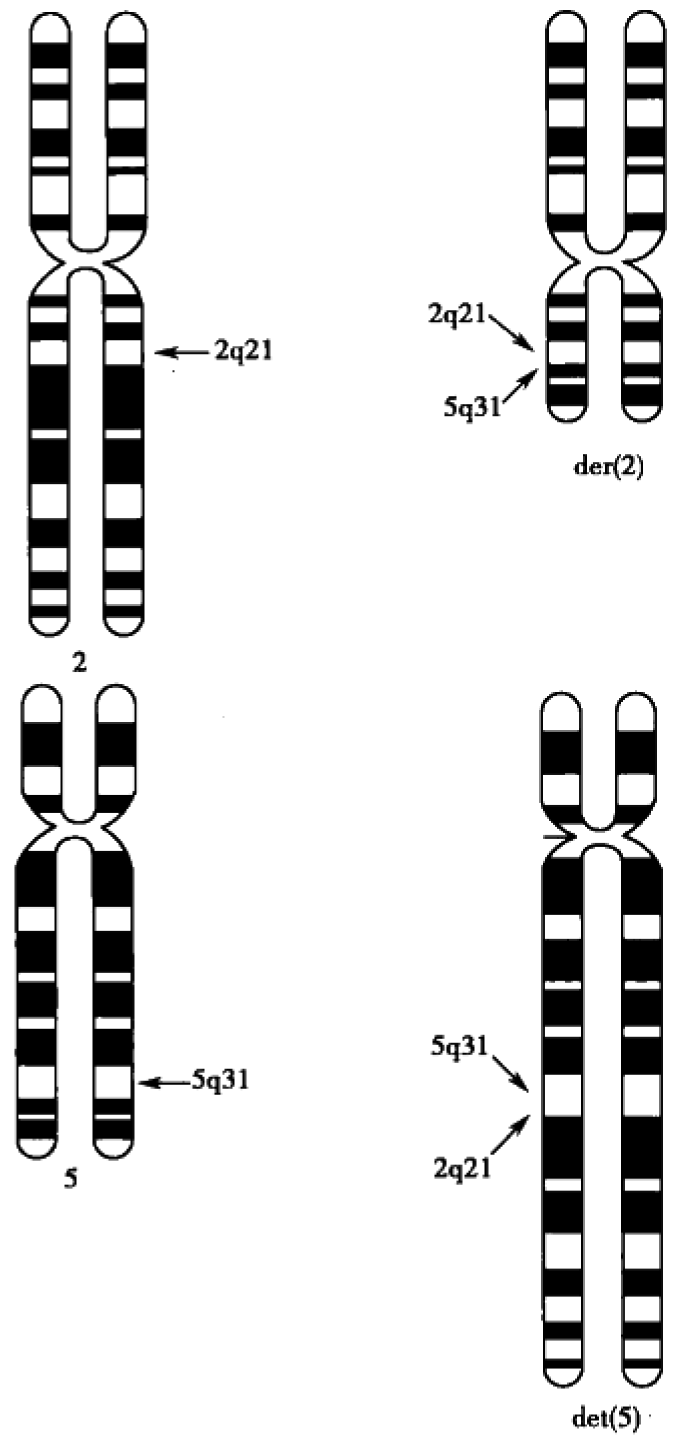


图4-18　染色体相互易位

5. 环状染色体（ring chromosome）　一条染色体的长、短臂同时发生了断裂，含有着丝粒的片段两断端发生重接，即形成环状染色体。如2号染色体的p21和q31分别发生了断裂，断点以远的片段丢失，含有着丝粒的中间片段两断端p21与q31相接形成环状染色体（图4-21）。这种结构畸变的简式描述为：46，XX（XY），r（2）（p21q31）；详式描述为：46，XX（XY），r（2）（p21→q31）。

6. 双着丝粒染色体（dicentric chromosome）：两条染色体同时发生一次断裂后，具有着丝粒的片段的两个断端相连接，形成了一条双着丝粒染色体。如6号染色体的q22和11号染色体的p15分别发生了断裂，两个具有着丝粒的染色体片段断端相互连接，形成了一条双着丝粒的衍生染色体（图4-

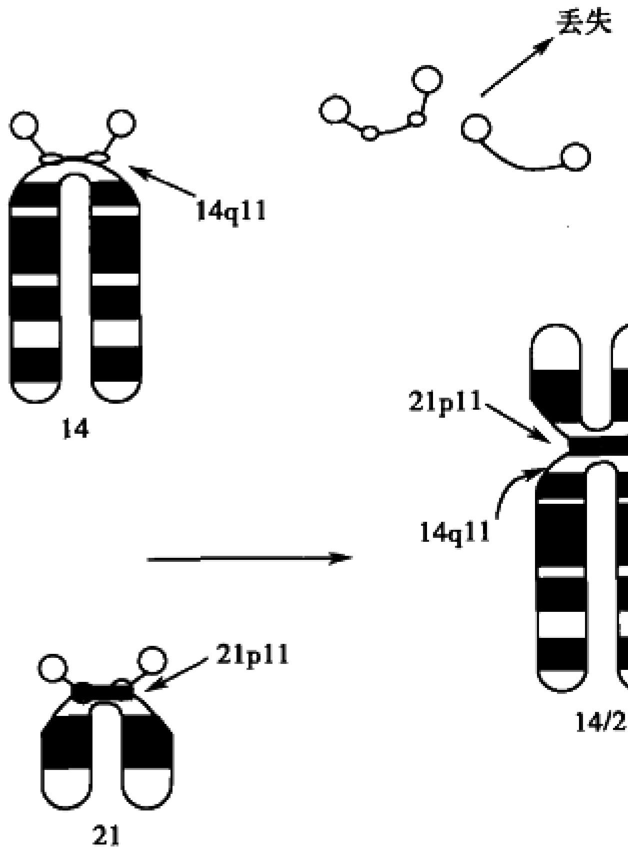


图 4-19 罗伯逊易位

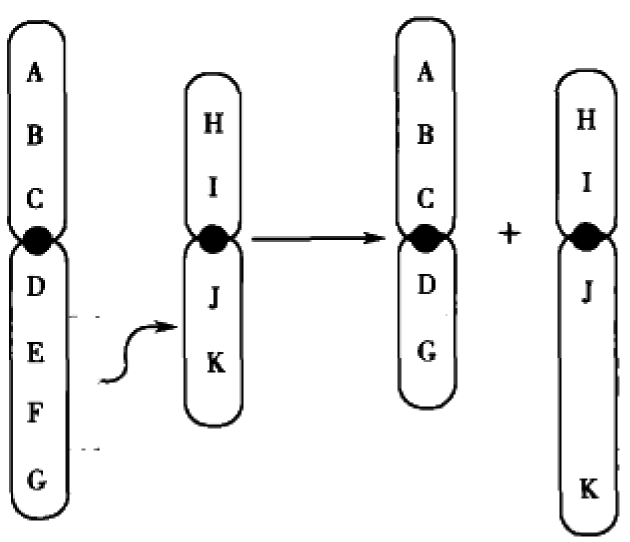


图 4-20 插入易位图解

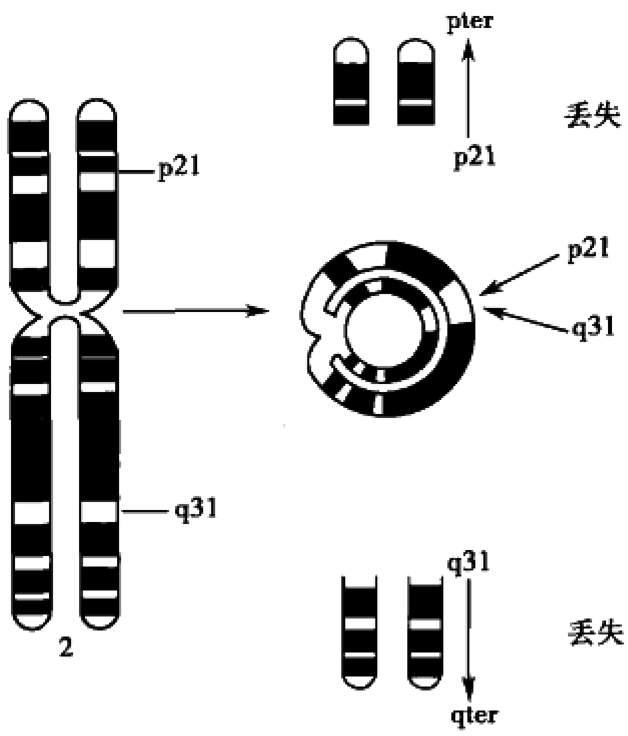


图 4-21 环状染色体

22)。这种结构畸变的简式描述为:45,XX,dic(6;11)(q22;p15);详式描述为:45,XX,dic(6;11)(6pter→6q22::11p15→11qter)。

7. 等臂染色体(isochromosome):一条染色体的两个臂在形态和遗传结构上完全相同,称为等臂染色体。等臂染色体一般是由于着丝粒分裂异常造成的。在正常的细胞分裂中,着丝粒纵裂,姐妹染色单体分离,形成两条具有长、短臂的染色体。如果着丝粒横裂,长臂、短臂各自形成一条染色体,即形成了一条具有两个长臂和一条具有两个短臂的等臂染色体。如图 4-23 所示。①具有两个长臂的等臂染色体的简式描述为:46,X,i(Xq);详式描述为:46,X,i(X)(qter→cen→qter);②具有两个短臂的等臂染色体的简式描述为:46,X,i(Xp);详式描述为:46,X,i(X)(pter→cen→pter)。

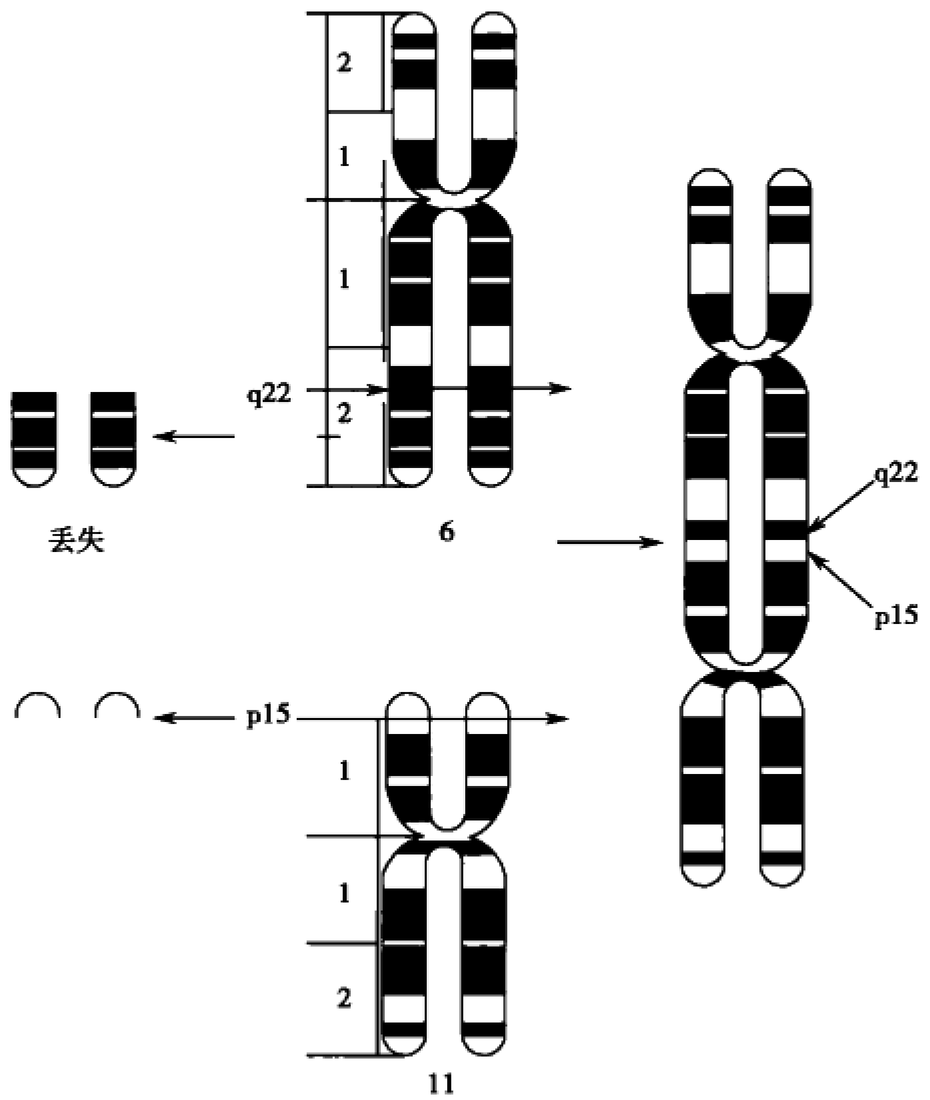


图 4-22　双着丝粒染色体

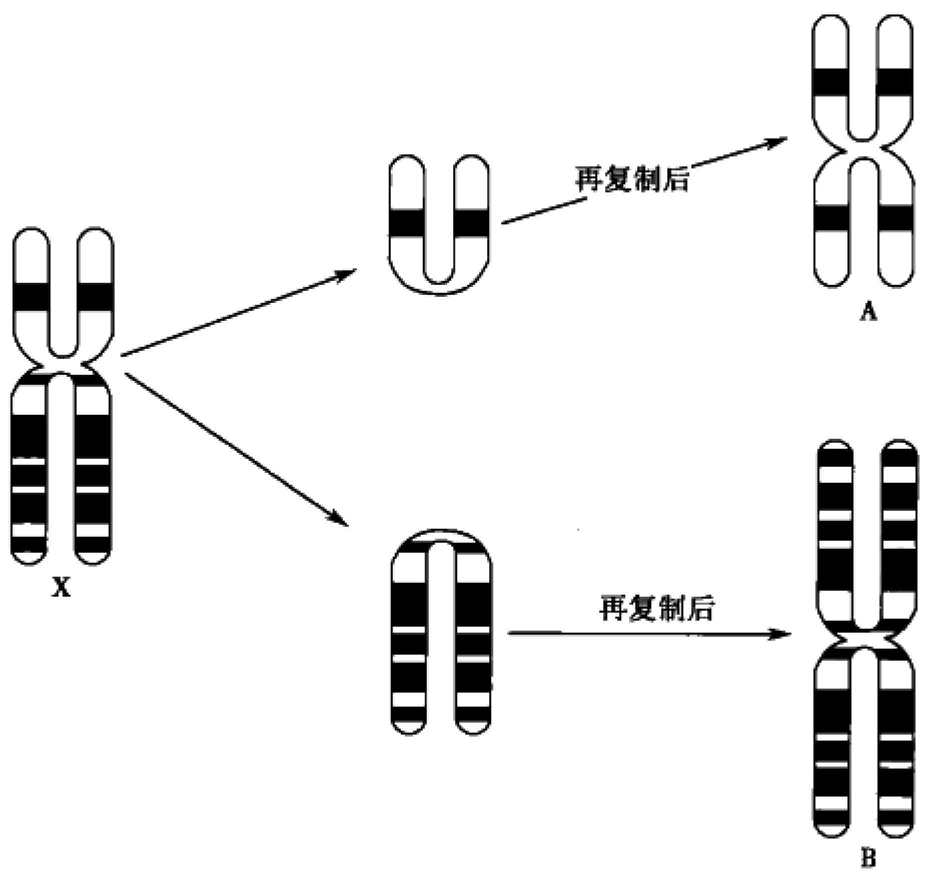


图 4-23　等臂染色体

第三节　染 色 体 病

染色体病(chromosome diseases)是染色体异常或畸变所致。由于染色体异常涉及许多基因,患者均有较严重或明显的临床症状,故又称染色体异常综合征。因此,染色体病一般具有以下临床特征:①染色体病患者一般有先天性多发畸形,智力发育和生长发育迟缓,有的还有特殊的皮肤纹理改变。具有染色体异常的胚胎,大部分流产或死产;②性染色体异常患者,除有上述特征外,还有内外生殖器异常或畸形,如性腺发育不良,副性征不发育等。

一、常染色体病

常染色体病(autosomal disease)是由常染色体数目或结构异常引起的疾病。常染色体病约占染色体病的2/3。包括三体综合征、单体综合征、部分三体综合征、部分单体综合征和嵌合体等。下面列举几种临床上较常见的常染色体病。

(一) Down 综合征(21 三体综合征)(OMIM190685)

Down 综合征(Down Syndrome,DS)也称 21 三体综合征(trisomy 21 syndrome)或先天愚型,1866年英国医师 John Langdon Down 首先对此病作了临床描述,故命名为 Down 综合征,长久以来不明其病因。但本病具有母亲生育年龄偏大和单卵双生的一致性两个特点很早就引起了一些人类遗传学家的注意。1932 年 Wardenburg 曾提议用染色体异常解释本病,但由于当时还没有合适的方法对他的这一论点加以验证。1959 年法国细胞遗传学家 Lejeune 等分析了 9 例先天愚型患儿的成纤维细胞的染色体,首先证实本病的病因是多了一个 G 组染色体(后来确定为 21 号),故本病又称为 21 三体综合征。21 三体综合征是发现最早的由于染色体异常而导致的疾病,也是最常见的染色体病。

1. Down 综合征的发病率　据统计新生儿中 Down 综合征的发病率约 1/800,我国目前大约有 60 万以上的 Down 综合征患儿,全国每年出生的 Down 综合征患儿可高达 27 000 例左右。Down 综合征发病率随母亲生育年龄的增高而增高,尤其当母亲年龄大于 35 岁时,发病率明显增高。

2. Down 综合征的临床特征　这一疾病的主要临床表现为生长发育迟缓,不同程度的智力低下和包括头面部特征在内的一系列异常体征。智力发育不全是本病最突出的症状,患者智商在 25 ~ 50。患者呈现特殊面容:眼距过宽、眼裂狭小、外眼角上倾、内眦赘皮、鼻梁低平、外耳小、耳廓常低位或畸形、硬腭窄小、舌大外伸、流涎,故又被称为伸舌样痴呆(图 4-24)。患者其他症状或体征还有:肌张力低下、四肢短小、手短宽而肥、第五手指因中间指骨发育不良而只有一条指横褶纹、肤纹异常,40% 有先天性心脏病,白血病的发病风险是正常人的 15 到 20 倍;患者 IgE 水平较低,容易发生呼吸道感染;白内障发病率较高。存活至 35 岁以上的患者出现老年性痴呆(Alzheimer 病)的病理表现。男性患者常有隐睾,无生育能力;女性患者通常无月经,偶有生育能力,并有可能将此病遗传给下一代。



图 4-24　Down 综合征患者

3. Down 综合征的遗传学类型　根据患者的核型组成不同,可将 Down 综合征分为三种遗传学类型。

(1) 21 三体型:也称游离型,具有三条独立存在的 21 号染色体。约占全部患者的 92.5%。核型为 47,XX(XY),+21(图 4-25)。

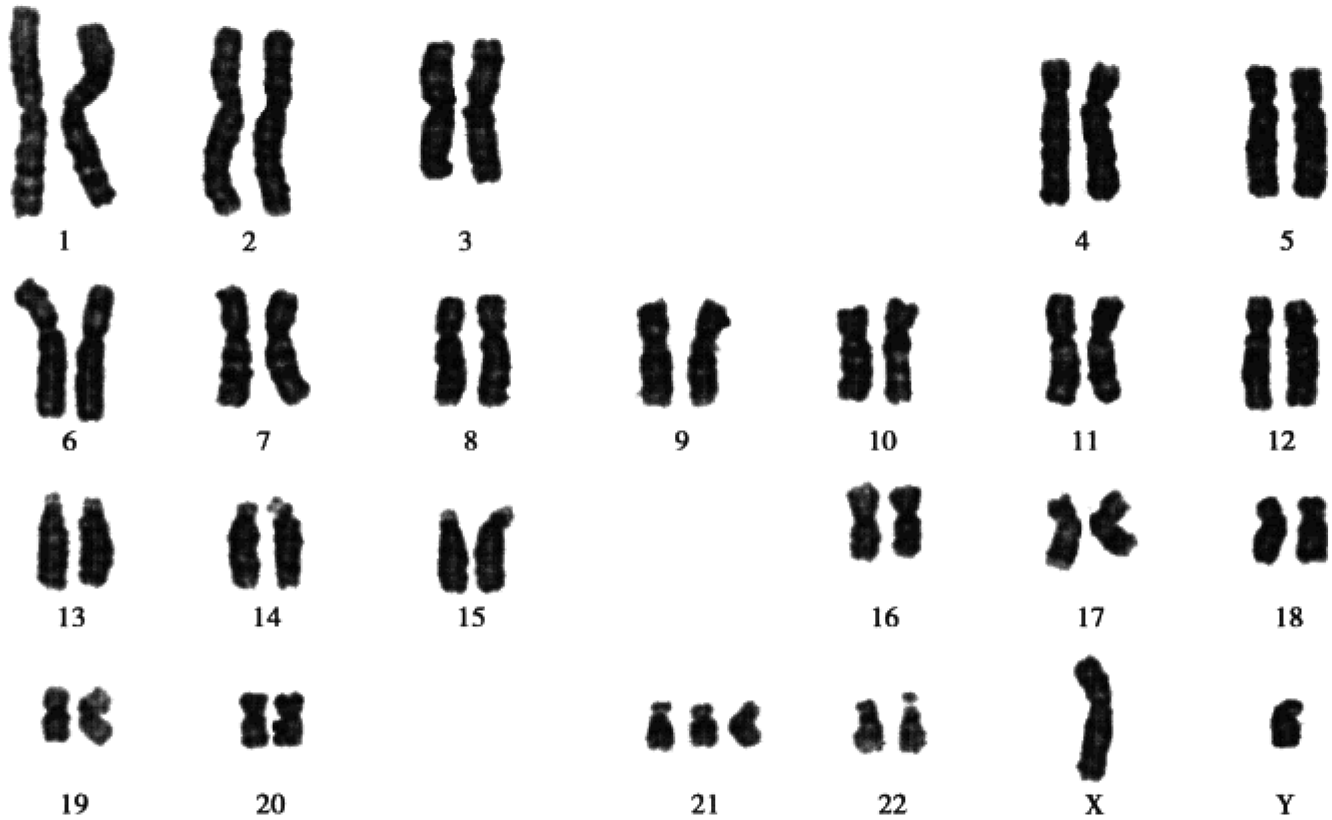


图 4-25 Down 综合征核型

此型的发生绝大部分与父母核型无关,它是生殖细胞形成过程中,在减数分裂时 21 号染色体发生不分离,结果形成染色体数目异常的配子,当其与正常的配子结合后,即产生 21 三体型的患儿。染色体不分离发生在母方的病例约占 95%,另 5% 见于父方,且主要为第一次减数分裂不分离。减数分裂不分离的机制还有待进一步研究,有研究表明可能与染色体支架蛋白-拓扑异构酶Ⅱ(topoⅡ)的活性改变有一定关系。此型的发生率随母亲的生育年龄增高而增高。生过此型患儿的父母,再生同类患儿的危险率为 1% ~2%。

(2) 易位型:此型约占全部患者的 5%。1960 年 Polani 首次报道了易位型先天愚型的病例。易位型患者具有典型的先天愚型临床症状。但其增加的一条 21 号染色体并不独立存在,而是与 D 组或 G 组的一条染色体发生罗伯逊易位,染色体总数为 46,其中一条是易位染色体,故称为假二倍体(pseudodiploid)。最常见的是 D/G 易位,如 14/21 易位,核型为 46,XX(XY), -14, +t(14q21q),其次为 G/G 易位,如 21/21 易位,核型为 46,XX(XY), -21, +t(21q21q)。患者的易位染色体,如果是由亲代传递而来的,其双亲之一通常是表型正常的染色体平衡易位携带者(balanced translocation carrier),其核型常为 45,XX(XY), -14, -21, +t(14q21q)。染色体平衡易位携带者在生殖细胞形成时,理论上经减数分裂可以产生 6 种类型的配子(见图 4-26),但实际上只有 4 种配子形成,故与正常个体婚配后,将产生 4 种核型的个体。由此可见,染色体平衡易位携带者虽外表正常,但其结婚怀孕后,常有自然流产或死胎史,所生子女中,约 1/3 正常,1/3 为易位型先天愚型患儿,1/3 为平衡易位携带者。但如果是 21/21 平衡易位携带者,即其核型为 45,XX(XY), -21, -21, +t(21q21q)者,其婚后所孕胎儿中,1/2 将因核型为 21 单体而流产,1/2 核型为 46,XX(XY), -21, +t(21q21q)。因此,活婴将 100% 为 21/21 易位型先天愚型患儿。对这种类型的携带者应告知其生育风险。易位型先天愚型一般常见于年龄较轻的父母所生子女。由于其双亲之一是染色体平衡易位携带者,故发病具有明显的家族倾向。

(3) 嵌合型:此型较少见,约占 2.5%。嵌合型产生的原因主要是由于受精卵在胚胎发育早期的卵裂过程中,第 21 号染色体发生不分离。如果第一次卵裂时发生不分离,就会产生 47, +21 和 45, -21 两个细胞系,而后一种细胞很难存活。因此,导致嵌合体的不分离多半发生在第一次卵裂以后的某次有丝分裂,形成 45/46/47 细胞系的嵌合体。所有嵌合体内都有正常的细胞系。不分离发生得越晚,正常细胞系所占比例就越多,则此患者症状就越轻。因本型患者的体细胞中含有正常细胞系,

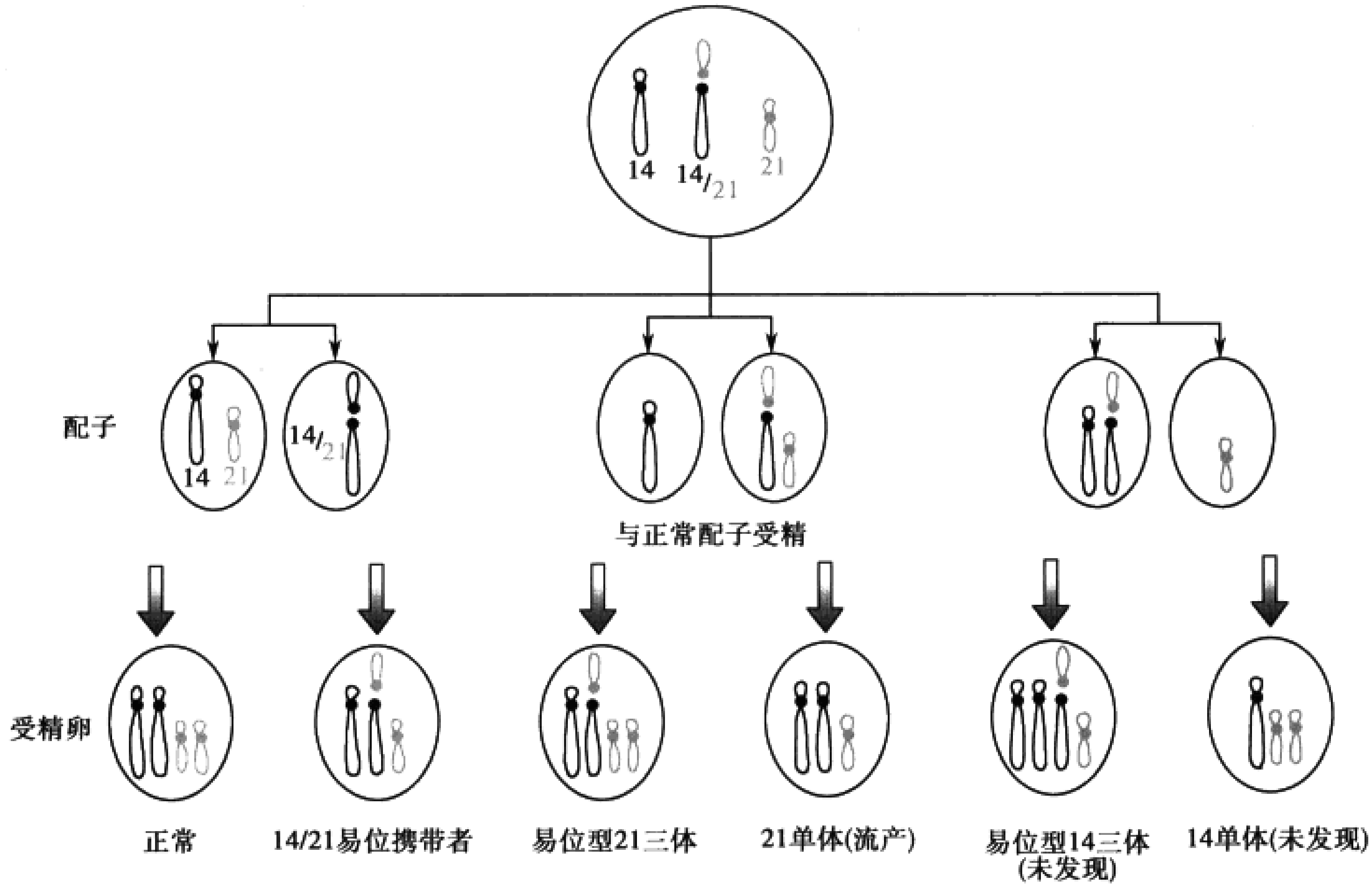


图 4-26　14/21 平衡易位携带者减数分裂后形成 6 种可能的配子及其后代核型图解

故临床症状多数不如 21 三体型严重、典型。

4. Down 综合征发生的分子机制　2000 年 5 月，由日、德等国科学家通力合作的人类 21 号染色体 DNA 序列测定工作完成。21 号染色体是人类染色体中最小的一条，由 5.1×10^7bp 组成，约长 46cM，包含 600～1000 个基因，占整个人类基因组的 1.7%（表 4-4）。

表 4-4　21q 各区带特定标记与相关表型

染色体分带	特定标记	相关的 21 三体表型
q11.1	D21S16 D21S13 D21S4	
q21	D21S52 D21S59 D21S1 D21S11 D21S8 D21S18 APP D21S54	智力发育迟缓（次要作用）
q 22.1	D21S93 SOD1 D21S82 D21S58 D21S65 D21S17	
q22.2	D21S55	智力发育迟缓（主要作用）
		肌张力低下、关节松弛、身材矮小
		9 种外貌特征（面、手、足）
	D21S3 HMG14	6 种外貌特征（面、皮纹）
q22.3	ETS2 D21S15 MX-1/2	先天性心脏疾患
	BCE1 D21S19 D21S42	
	CBS CRYA1	
	PFKL CD18 COL6A1/2 S100B	

通过对部分 21 三体的基因型与表型关系的研究，现已将 DS 的 24 种特征定位在 21 号染色体的 6 个小区域，其中 D21S52 和 D21S55-MX1 两个区域尤为引人关注。D21S52 是表达 13 种特征的最小区域，13 种特征分别是：智力障碍、身材矮小、肌张力下降、关节松弛和 9 种面貌特征：鼻梁扁平、舌外伸、腭弓高、窄腭、耳廓畸形、手掌宽且短、第五指短且弯、足第一、二趾间距宽。D21S55-MX1 是表达 6 种外貌特征（眼裂斜、内眦赘皮、Brushfield 斑-虹膜周围小白斑、通贯手、指纹尺箕和小鱼际肌无侧环）

的最小区域。D21S55 在 DS 的发病机制中起重要作用，在 21q22.2 跨 0.4 ~ 3kb。D21S55 及 21 q22.3 远端被称为 DS 关键区（Down's syndrome critical region，DCR）。一些研究已显示与 DS 发病有关的基因可能是一些结构基因或调控基因，但具体作用机制尚不太清楚。

21 号染色体 DNA 测序的完成，无疑将加快 21 号染色体基因功能的研究，对揭示 DS 及其他疾病的分子病因，更快、更精确地诊断 DS，并且在分子病理学上进行干预性治疗，都具有深刻的意义。

5. Down 综合征预后　先天愚型胎儿中，有 3/4 自发流产，且大部分发生在妊娠 3 个月内，仅约 1/4 能活到出生。患者智力低下，缺乏抽象思维能力，精神运动性发育缺陷，但许多患者经过训练可以学会读和写，以及一些基本的生活技能。一些人还可以达到接近边缘的社会适应力。但绝大部分患者都没有自理能力。随着医疗水平的不断提高，现在的先天愚型患者的生存期增加。许多人可以活到成年。但一般寿命比正常人短，只有 8% 的患者活过 40 岁。

（二）18 三体综合征

1960 年 Edwards 等首先报告本病，故又称为 Edwards 综合征（Edwards syndrome）。1961 年 Patau 证实了该症的病因是多了一条 18 号染色体，因此定名为 18 三体综合征（trisomy 18 syndrome）。

1. 18 三体综合征的临床特征　新生儿发病率约为 1/3500 ~ 1/8000，但在某些地区或季节明显增高，达到 1/800 ~ 1/450。男女性别比为 1:4，这可能与此类男性胚胎不易发育至出生有关。患者宫内生长迟缓，小胎盘及单一脐动脉，胎动少，羊水过多，95% 胎儿流产；一般过期产，平均妊娠 42 周；出生时体重低，平均仅 2243 克，发育如早产儿，吸吮差，反应弱，因严重畸形，出生后 1/3 在 1 个月内死亡，50% 在 2 个月内死亡，90% 以上 1 岁内死亡，只有极个别患者活到儿童期。

18 三体综合征的主要临床特征为生命力严重低下，多发畸形，生长、运动和智力发育迟缓。其异常表型主要有：眼裂小、眼球小、内眦赘皮、耳畸形伴低位、枕骨突出、小颌、唇裂或腭裂、胸骨小；95% 有先天性心脏病，它构成了婴儿死亡的主要原因；手呈特殊握拳姿势：第 2 和第 5 指压在第 3 和第 4 指之上；有所谓"摇椅样畸形足"（图 4-27）。

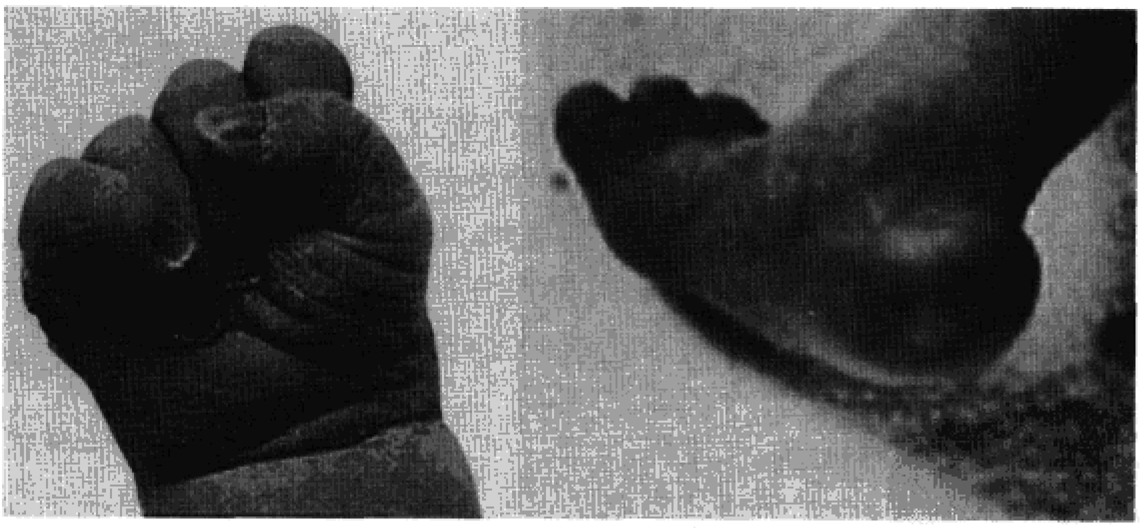


图 4-27　18 三体综合征患者的特殊握拳姿态和摇椅形足

2. 核型与遗传学　本症患者有 80% 核型为 47，XX（XY），+18，症状典型（图 4-28）。18 三体型的产生多由母亲卵母细胞减数分裂时发生的 18 号染色体不分离所致，其发生与母亲年龄增大有关；10% 为嵌合型，即 46，XX（XY）/47，XX（XY），+18，症状较轻；其余为各种易位，主要是 18 号与 D 组染色体易位。

（三）13 三体综合征

1960 年由 Patau 首先描述的一个具有一条额外的 D 组染色体的患儿，故又称为 Patau 综合征。后来通过显带技术确定多余的染色体为 13 号染色体。

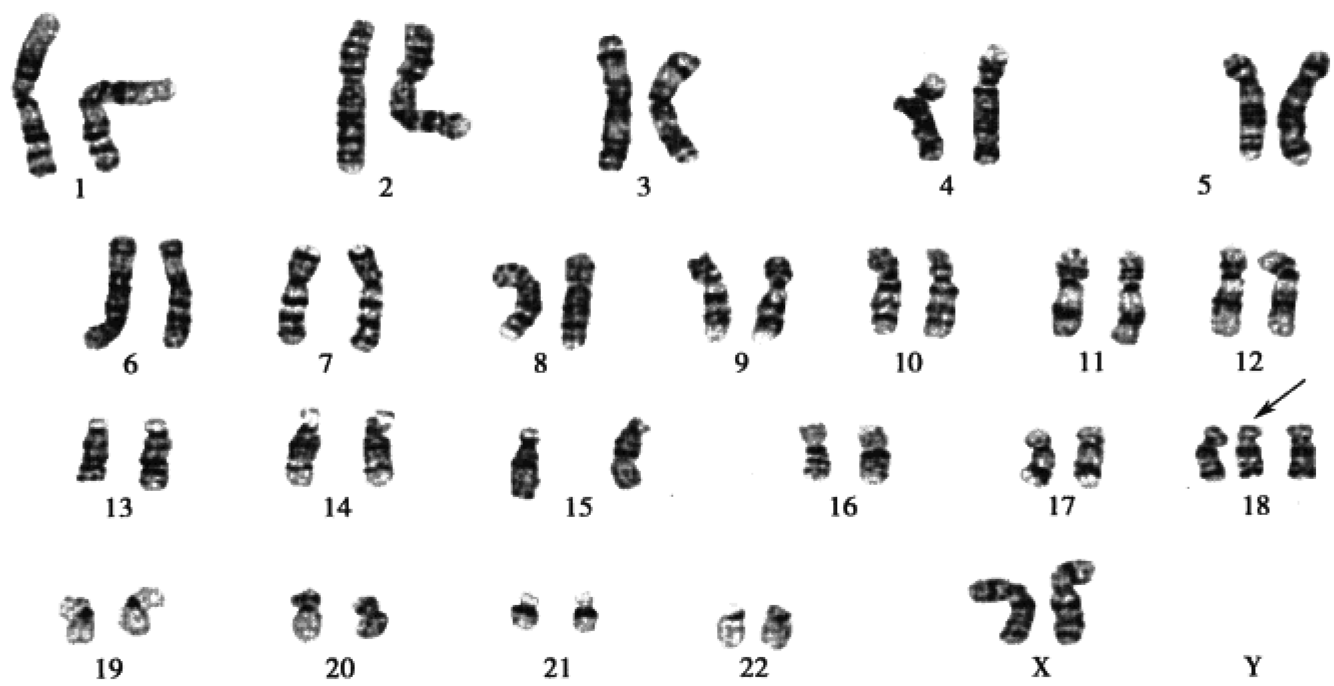


图 4-28 18 三体综合征核型

1. 13 三体综合征的发病率及临床特征 新生儿中 13 三体综合征的发病率约为 1/25 000，女性明显多于男性。发病率与母亲年龄增大有关。患者的畸形比上述两种综合征严重。99% 的 13 三体型胚胎导致流产。

13 三体综合征的主要临床特征是中枢神经系统发育严重缺陷，无嗅脑，前脑皮质形成缺如，称为前脑无裂畸形；出生体重低、发育迟缓、严重智力低下、小头、小眼球或无眼球、小颌、多数有唇裂或伴腭裂、耳低位畸形、常有耳聋，80% 有先天性心脏病，1/3 有多囊肾，无脾或有副脾，男性有隐睾，女性多有双角子宫及卵巢发育不全，常有多指，有与 18 三体综合征相似的特殊握拳姿势和摇椅样畸形足、肤纹异常等（图 4-29）。



图 4-29 13 三体综合征患者

2. 核型与遗传学 患者中 80% 的核型为 47，XX（XY），+13（图 4-30），其发生与母亲年龄有关，额外的 13 号染色体大多来自母方减数分裂 I 的染色体不分离。10% ~15% 为易位型，多为 13/14 的罗伯逊易位，易位型多为年轻母亲所生，她们常有流产史。5% 是嵌合型，即核型为 46，XX（XY）/47，XX（XY），+13，一般症状较轻。

（四）5p⁻综合征（OMIM123450）

1963 年由 Lejeune 等首先报道，因患儿具特有的猫叫样哭声，故又称为猫叫综合征（cri du chat syndrome）。1964 年证实本病症为第 5 号染色体短臂部分缺失所致，故也称为 $5p^-$ 综合征。

1. $5p^-$ 综合征的发病率及临床特征 该综合征发病率在新生儿中为 1/50 000，在智能低下患儿中约占 1% ~1.5%。本病的最主要临床特征是患儿在婴幼儿期的哭声似猫叫。其他症状有生长、智力发育迟缓、小头、满月脸、眼距较宽、外眼角下斜、斜视、内眦赘皮、耳低位、小颌、并指、髋关节脱臼、肤纹异常、50% 有先天性心脏病等。多数患者可活至儿童期，少数活至成年，均伴有严重智力低下。在智商低于 35 的群体中约占 1%。

2. 核型与遗传学 核型为 46，XX（XY），$5p^-$（图 4-31）；也有部分是嵌合型。患者 5 号染色体短臂缺失的片段大小不一，经多个 DNA 探针检测，证实缺失片段为 5p15，即本病是 5p15 缺失引起。80% 的病例为染色体片段的单纯缺失（包括中间缺失），10% 为不平衡易位引起，环状染色体或嵌合



图 4-30　13 三体核型

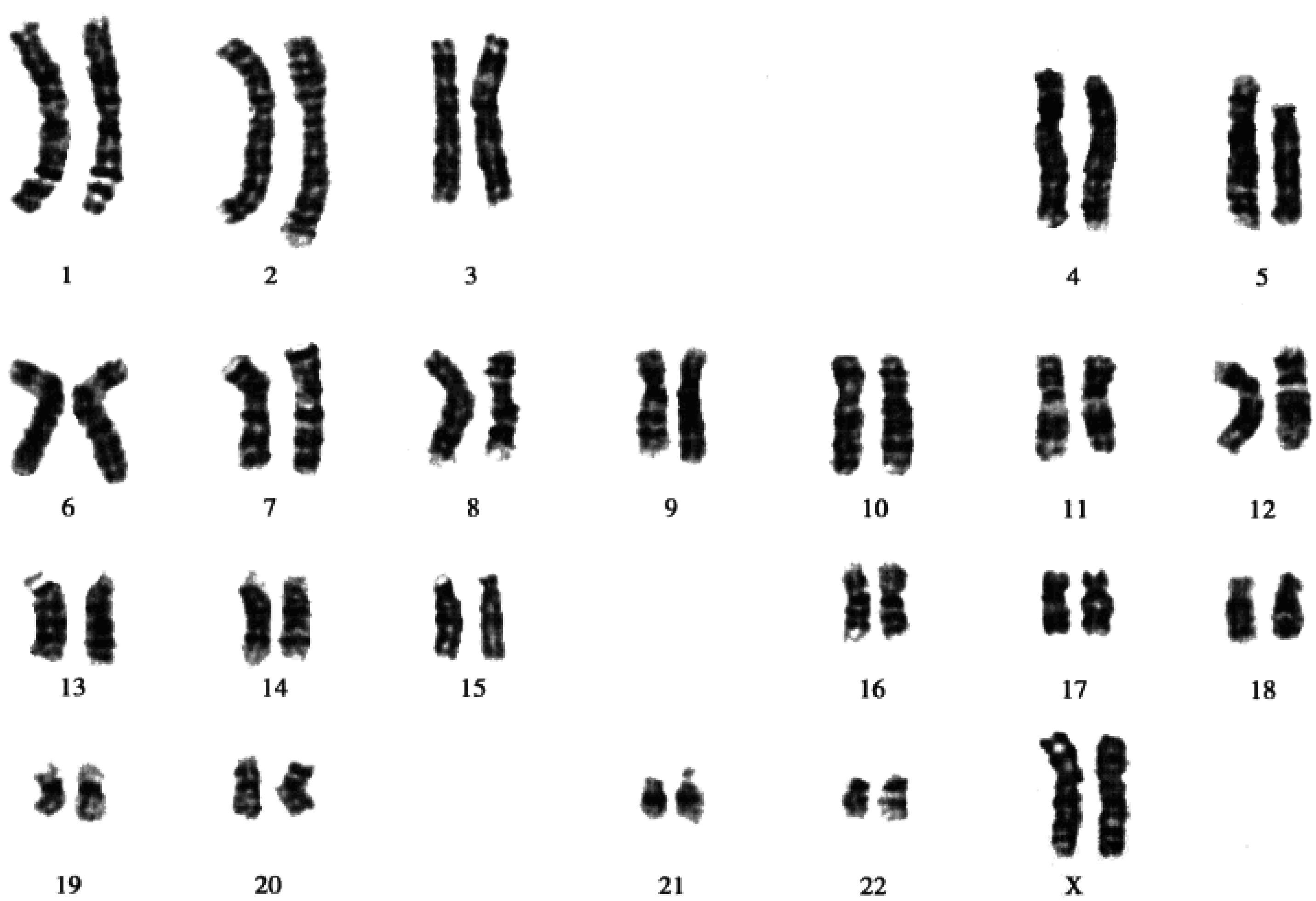


图 4-31　猫叫综合征患者核型图

体则比较少见。多数病例是父母生殖细胞中新发生的染色体结构畸变所引起，约有 10% ~15% 是平衡易位携带者产生的异常配子所引起。

二、性染色体病

性染色体病(sex chromosome disease)指性染色体 X 或 Y 发生数目或结构异常所引起的疾病。性染色体虽然只有一对，但性染色体病约占染色体病的 1/3。临床上较常见的性染色体病为：

(一) Klinefelter 综合征

1942 年 Klinefelter 等首先报道了该综合征，故称为 Klinefelter 综合征，也称先天性睾丸发育不全或原发性小睾丸症。1956 年 Bradbury 等在患者的细胞内发现 X 染色质阳性(正常男性 X 染色质为

阴性),1959 年 Jacob 和 Strong 证实患者的核型为 47,XXY,即较正常男性多出一条 X 染色体,因此又叫做 47,XXY 综合征。

1. Klinefelter 综合征的发病率及临床特征　本病发病率相当高,在男性新生儿中占 1/1000 ~ 2/1000,在身高 180cm 以上的男性中占 1/260,在精神病患者或刑事收容所中占 1/100,在不育的男性中占 1/10。

Klinefelter 综合征以睾丸发育障碍和不育为主要特征。第二性征发育不良,阴茎发育不良、睾丸小或隐睾,精曲小管萎缩并呈玻璃样变性,不能产生精子,因而不育。患者体征呈女性化倾向,大部分人无胡须、无喉结、体毛稀少、阴毛呈女性分布、稀少或无毛、皮下脂肪丰富、皮肤细嫩、约 25% 的个体发育出女性型乳房、其性情和体态趋向于女性特点(图 4-32)。此外还可能有头围小、指距宽、耳畸形、骨骼异常、先天性心脏病等畸形;部分患者有轻度到中度智力障碍,表现为语言能力低下,一些患者有精神分裂症倾向。

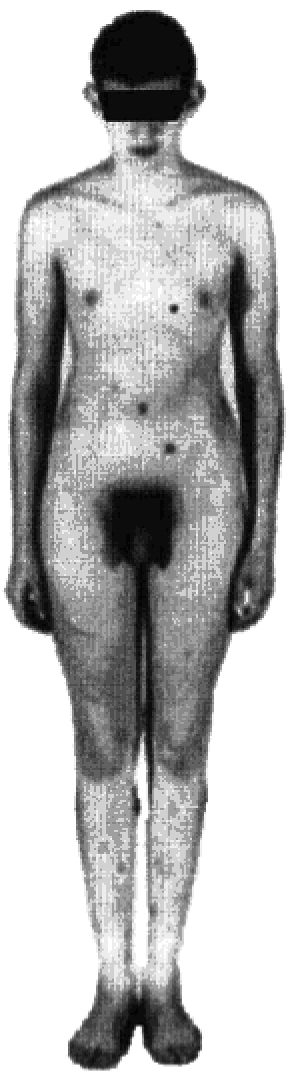


图 4-32　Klinefelter 综合征患者

2. 核型与遗传学　患者的主要核型为 47,XXY,约占 80%,见图 4-33。嵌合型约占 15%,包括 46,XY/47,XXY;45,X/46,XY/47,XXY;46,XX/47,XXY 等。另外还可见核型如 48,XXXY,48,XXYY,49,XXXXY 等。一般来讲,核型中 X 染色体数量越多,表现的症状越严重。例如 49,XXXXY 的个体除了上述症状更明显外,还有智力极度低下,并具有小头、颈蹼、腭裂、桡尺骨连合、肘外翻、膝外翻、脊柱畸形等异常。而嵌合型的症状相对较轻且不典型。本征额外的 X 染色体产生于减数分裂时 X 染色体的不分离,不分离发生在父方和母方的几率均等。

对本综合征患者可在青春期用雄激素替代治疗,以维持男性表型,改善患者心理状态。但如果疗效不佳,应停止使用激素。男性乳房发育,可手术切除。凡具有 Y 染色体而性腺发育不良者,易有性腺恶变,应给予重视。

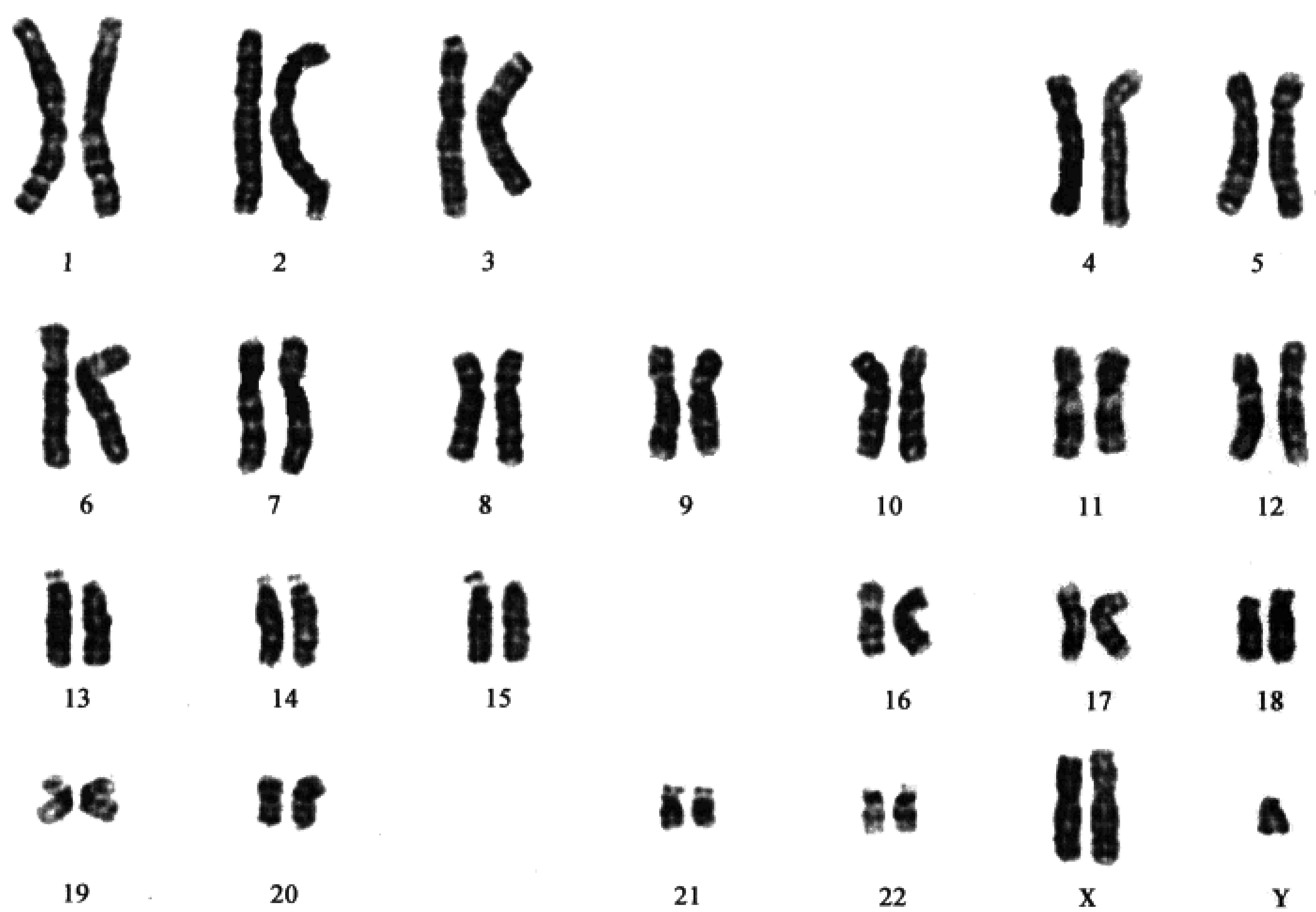


图 4-33　Klinefelter 综合征核型

（二）Turner 综合征

1938 年 Turner 首先描述该综合征，故命名为 Turner 综合征，随后发现患者体内有条索状卵巢，无卵泡发生，因此称为性腺发育不全或先天性卵巢发育不全。1954 年发现多数患者的 X 染色质阴性。1959 年 Ford 等证实患者的核型为 45，X。

1. Turner 综合征的发病率及临床表现　在新生女婴中 Turner 综合征的发病率约为 1/5000，但在自发流产胎儿中可高达 18% ~20%。本病在胎儿中占 1.4%，但在子宫内不易存活，其中 99% 流产。

本综合征的主要临床特征是，表型女性；有出生体重低，新生儿期脚背有淋巴样肿，第 4、第 5 指骨短小或畸形；身材发育缓慢尤其缺乏青春期发育，使成年身材显著矮小，仅在 120 至 140cm 之间；后发际低，头发可一直延至肩部；50% 个体出现颈蹼；还可有盾状胸、肘外翻、两乳头间距过宽、肤纹异常等。第二性征发育差，表现为成年外阴幼稚、阴毛稀少、乳房不发育、子宫发育不良、卵巢无卵泡、原发闭经，因而不能生育（图 4-34）。此外，约 1/2 患者有主动脉狭窄和马蹄肾等畸形。部分患者有轻度到中度智力障碍，表现为语言能力低下，一些患者有精神分裂症倾向。

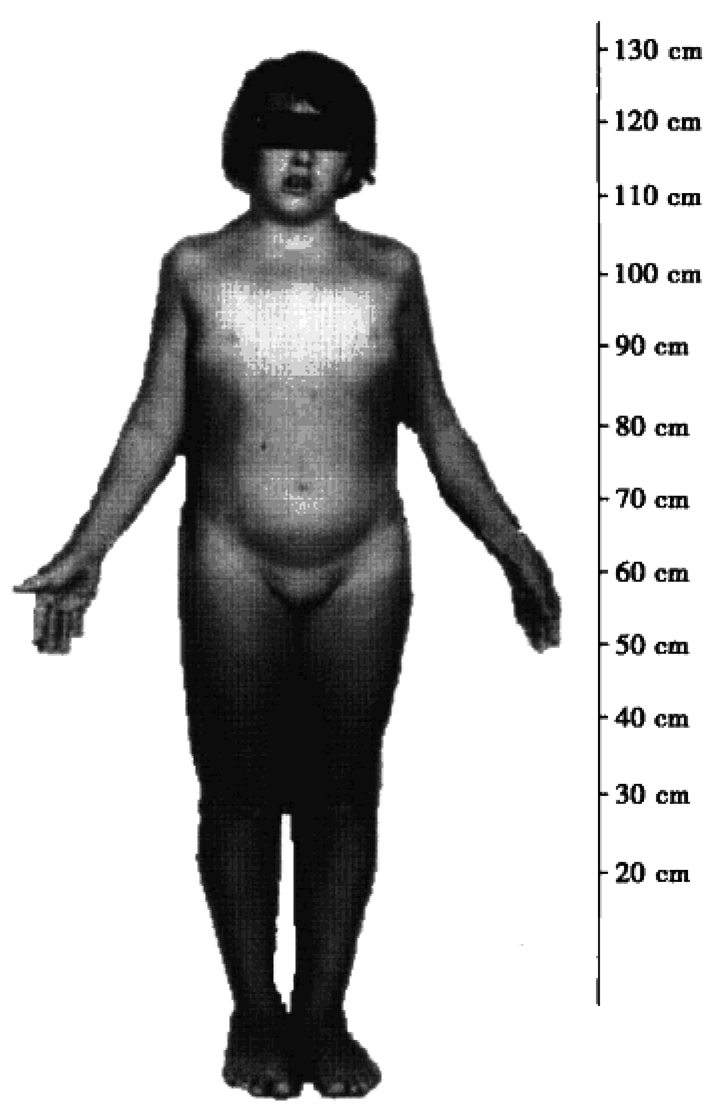


图 4-34　Turner 综合征患者

2. 核型　约 55% 病例核型为 45，X（图 4-35），还有各种嵌合型和结构异常的核型，最常见的嵌合型为 45，X/46，XX。结构异常有 X 等臂染色体，其核型为 46，X，i（Xq）。一般来说，嵌合型的临床表现较轻，轻者有可能有生育力，而有 Y 染色体的嵌合型可表现出男性化的特征；身材矮小和其他 Turner 体征主要是由 X 短臂单体性决定的；但卵巢发育不全和不育则更多与长臂单体性有关。

除少数患者由于严重畸形在新生儿期死亡外，一般均能存活。对性腺发育不全的治疗原则主要是对症治疗。在青春期用女性激素治疗可以促进第二性征和生殖器官的发育，月经来潮，改善患者的心理状态，但不能促进长高和解决生育问题。对大量患者的治疗尚在研究中。

（三）XYY 综合征

1961 年由 Sandburg 等首次报告。本病在男性中的发生率为 1/900。核型为 47，XYY，额外的 Y 染色体肯定来自父方精子形成过程中减数分裂Ⅱ时发生 Y 染色体的不分离。XYY 男性的表型一般正常，患者身材高大，常超过 180cm，偶尔可见尿道下裂，隐睾，睾丸发育不全并有生精过程障碍和生育力下降；但大多数男性可以生育，个别患者生育 XYY 的后代。

47，XYY 的核型中额外的 Y 染色体来源于父亲 Y 染色体减数分裂不分离，但也有来自 47，XYY 父亲的生殖细胞发生的次级不分离。此外，少数个体还有 48，XXYY；49，XYYYY；48，XYYY；47，XYY/46，XY；45，X/49，XYYYY 等特殊核型。此时 Y 染色质的检查会显现出相应数量的 Y 荧光小体。一般来讲，核型中 Y 染色体越多，这些类型的患者会出现智力发育的严重障碍和各种严重畸形。

（四）X 三体综合征

1959 年由 Jacob 首先报道。本病发病率在新生女婴中为 1/1000。在女性精神病患者中发病率约为 4/1 000。X 三体女性可无明显异常，约 70% 病例的青春期第二性征发育正常，并可生育；另外 30% 患者的卵巢功能低下，原发或继发闭经，过早绝经，乳房发育不良；1/3 患者可伴先天畸形，如先

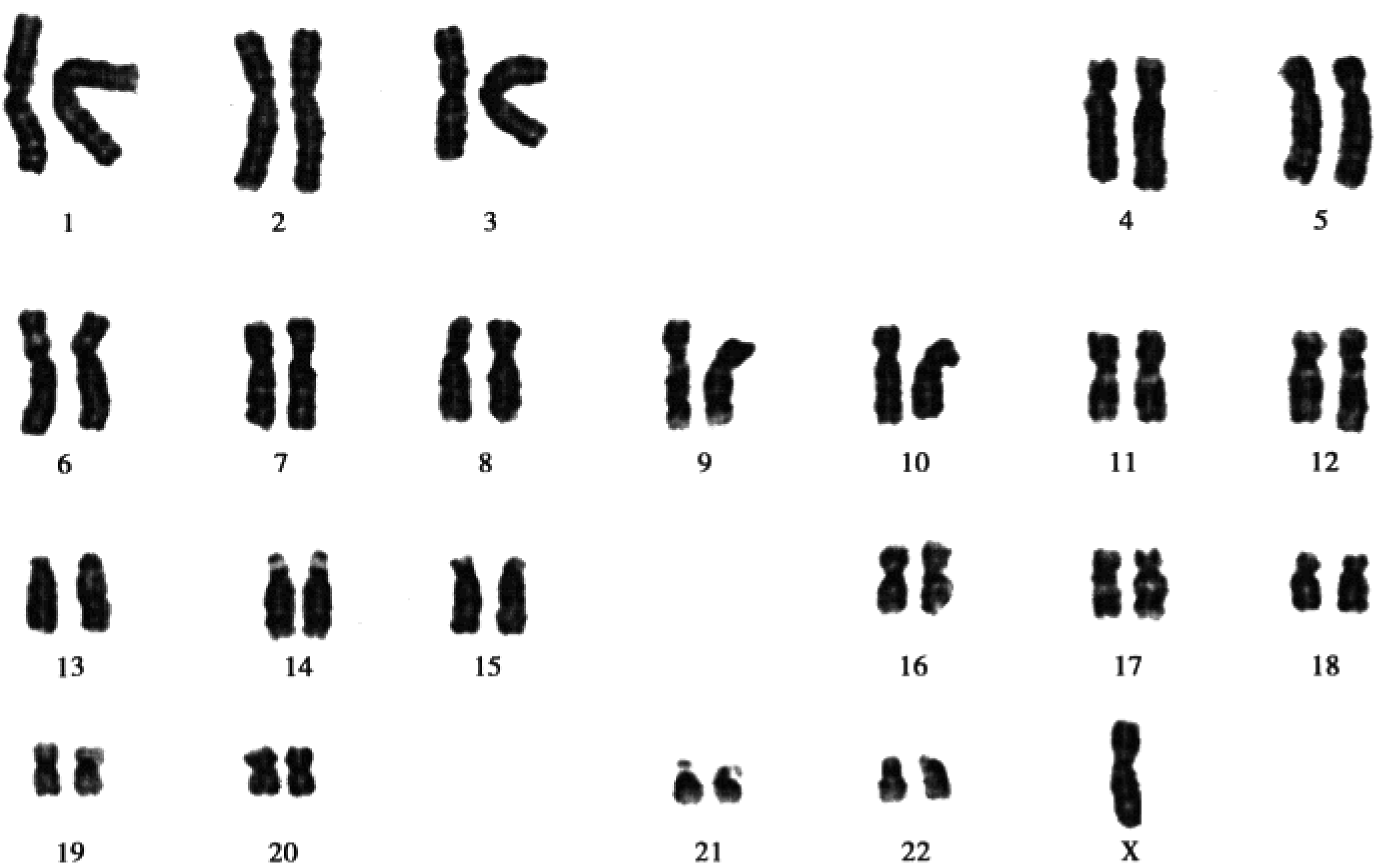


图 4-35 Turner 综合征核型图

天性心脏病、髋脱位；部分可有精神缺陷。约 2/3 患者智力低下。X 染色体越多，智力发育越迟缓，畸形亦越多见。患者核型多为 47，XXX。体细胞中有两个 X 染色质。少数核型为 46，XX/47，XXX。极少数为 48，XXXX、49，XXXXX。额外的 X 染色体，几乎都来自母方减数分裂的不分离，且主要在减数分裂 I 时发生。

（五）脆性 X 染色体综合征（fragile X syndrome）

当外周血淋巴细胞在缺乏叶酸或胸腺嘧啶的培养基中培养后，其染色体上就可以观察到明显的断裂或裂隙，这些断裂或裂隙称为脆性部位。脆性 X 染色体是指在 Xq27.3 位置具有脆性部位的 X 染色体。

1. 脆性 X 染色体综合征的临床症状　这一 X 染色体脆性部位是 1969 年由 Lubs 在一个 X 连锁的智力低下家庭中发现的，这个 X 连锁的智力低下就是后来知道的脆性 X 染色体综合征。脆性 X 染色体综合征的发病率在男性约为 1/1250，在女性约为 1/2000，没有明显的种族特异性。

脆性 X 染色体综合征的临床症状：受累男性表现为中度（IQ = 35 ~ 49）至重度（IQ = 20 ~ 34）智力低下，表现在语言障碍和算术能力差；还可表现为多动症、性情孤僻、精神病倾向。各种体征包括：大睾丸、大耳、长形面容、前额和下颌突出。其中巨大的睾丸是青春期以后出现的典型体征。但患者的睾丸功能正常，可有正常的生育能力。受累女性的临床表现通常较轻，1/3 的女性杂合子有轻度智力发育障碍，其发病与女性正常的 X 染色体随机失活，而脆性 X 染色体在众多体细胞中保持活性有关。但女性只有遗传自母亲携带者时才发病。正常男性携带者的女儿不发病，但外孙（女）可能发病。该病在连续遗传中有早现现象，即发病年龄有一代代提前并加重的倾向。

2. 脆性 X 染色体综合征的分子机制　与脆性 X 染色体智力低下有关的基因已被克隆，并被命名为 *FMR1*（Fragile X mental retardation 1）。该基因位于 Xq27.3，长 38kb，包含 17 个外显子，其表达最高的组织包括脑、睾丸及卵巢。该基因的 5′末端有一三聚核苷酸重复序列（CGG）n，CGG 重复序列的长短在人群中具有多态性，正常人可具有 6 ~ 50 个 CGG 重复序列。脆性 X 染色体智力低下患者具有 200 ~ 1000 个 CGG 重复序列。当重复次数达到约 200 次后，*FMR1* 基因的 5′端发生异常甲基化，导致基因转录失活而发病。CGG 串联重复次数的增加和相邻区域的高度甲基化也造成了脆性 X 染色体脆性部位的显示。当一个个体 CGG 的重复次数达到 52 次后，这一区域在减数分裂过程中即显现不

稳定状态，其重复次数可继续增加。重复次数在 52～200 次之间称为前突变(premutation)，带有前突变的个体称为携带者。前突变在遗传过程中不稳定，携带者在减数分裂过程中 CGG 串联重复继续增加至 200 次以上并使相邻区域高度甲基化，称为全突变(full mutation)，具有全突变的所有男性和约半数女性在临床上发病。但全突变只产生于前突变，不能由正常重复的 CGG 形成。而且携带者男性在生女儿时并不发生全突变。前突变携带者女性不表现症状，但在传给子代时重复序列进一步延长，达到全突变的长度，其子代出现症状。此外，CGG 发生前突变后在有丝分裂时也表现不稳定，因此受累个体的体细胞中可继续发生 CGG 不同次数的扩增，形成体细胞的"嵌合"性，即不同体细胞 CGG 的重复次数不同。这样一种基因突变的形式被称为"动态突变"。CGG 重复序列不稳定性和延长特征，可以解释为什么脆性 X 染色体智力低下综合征的遗传不遵循孟德尔规律。

目前已经发现类似的三核苷酸串联重复的动态增加也是许多其他单基因遗传病，例如 Huntinton 舞蹈病，强直性肌营养不良等的致病原因。

由于已经基本了解这一疾病的分子基础，所以分子诊断技术将比细胞遗传学分析更加有效和可靠。对产前或出生后个体的血液或组织样品提取 DNA，用两种限制性内切酶处理，其中一种不能切割甲基化的 DNA，这样就可对 DNA 进行甲基化分析并估计 CGG 串联重复的长短。另一方法是运用 PCR 技术判断 CGG 串联重复的次数。

（六）两性畸形

个体的性腺或内、外生殖器、第二性征具有不同程度的两性特征称为两性畸形(hermaphroditism)。两性畸形形成的原因很复杂，性染色体的畸变有时导致两性畸形的发生，但并非所有两性畸形都是由于性染色体异常引起。某些单基因的缺陷和环境因素也可造成两性畸形。内外生殖器和各种第二性征的发育主要受到性激素的作用。在性别分化和发育过程中，由于遗传或环境因素的影响使性激素的分泌或代谢发生紊乱，都可导致两性畸形。在一些遗传病中，两性畸形可作为多发畸形的体征之一而表现。这类患者可由基因突变所致，虽然染色体正常，但染色体检查可确定核型，将有助于诊断。

1. 真两性畸形　患者既有睾丸又有卵巢，内外生殖器间性，第二性征发育异常。核型：约 57% 为 46,XX；12% 为 46,XY；5% 为 46,XX/46,XY；其余为各种染色体异常。

2. 假两性畸形　假两性畸形不是真正的两性畸形，而是一种与核型相符的性腺。患者体内只存在一种性腺，但外生殖器和第二性征兼有两性特征，或者倾向于相反的性别。根据性腺为睾丸或卵巢，可将其分为男性假两性畸形和女性假两性畸形。

（1）男性假两性畸形：核型为 46,XY。性腺为睾丸，内外生殖器具有两性特征，第二性征异常。部分有女性化表型。

1）雄激素不敏感综合征(androgen insensitivity syndrome, AIS)：又称睾丸女性化综合征(testicular feminization syndrome)，为雄激素受体基因突变，呈 XR 遗传，患者外表可完全女性化或呈间性。有女性外阴，但无女性内生殖器，睾丸位于腹腔或腹股沟内，后者常被误认为是疝气。血中睾酮在正常水平。病因是 X 染色体上雄激素受体基因突变，致使靶细胞对雄激素不敏感。常因无月经或不孕而就诊。

2）特发性男性假两性畸形：为常染色体隐性遗传。患者体内雄激素合成不足而导致性发育异常。

3）Smith-Lemili-Opitz 综合征：为胆固醇合成酶缺陷，呈 AR 遗传，基因定位于 11q12.13。男性患者有隐睾、鼻短、鼻孔向前、腭裂、多指、骨骼异常、幽门狭窄等。1982 年 Lawry 发现本病在加拿大 British Columbia 地区为第二常见的隐性遗传病，仅次于囊性纤维变性(CF)，故应引起重视。

（2）女性假两性畸形：核型为 46,XX。性腺为卵巢，内外生殖器呈间性，第二性征发育有男性化倾向。常见有先天性肾上腺增生症(congenital adrenal hyperplasia, CAH)，呈 AR 遗传，基因定位于 6p21.3，其中以 21 羟化酶缺陷为多见，其次为 11 羟化酶缺陷，部分患者还伴有水盐代谢紊乱。

Summary

Chromosomes are complex structures located in the cell nucleus, they are composed of DNA, histone and non-histone proteins, RNA, and polysaccharides. They are basically the "packages" that contain the DNA. Each of the human chromosomes contains hundreds to thousands of genes, the major functional units of DNA.

Chromosome abnormalities may be either numerical or structural and may involve one or more autosomes, sex chromosomes or both simultaneously. Various human disorders result from abnormal chromosome number or structure. Such disorders often result in a syndrome, which is a group of symptoms that always occur together. They account for a large proportion of all reproductive wastage, congenital malformations, and mental retardation and play an important role in the pathogenesis of malignancy. Chromosome disorders are present in nearly 1 percent of live births, in about 2 percent of pregnancies in women older than 35 years who undergo prenatal diagnosis, and in fully half of all spontaneous first-trimester abortions.

(吴白燕)

参考文献

1. Gelehrter TD. 医学遗传学原理. 孙开来译. 北京:科学出版社,2001
2. 李璞. 医学遗传学. 第2版. 北京:中国协和医科大学出版社,2004
3. Strachan T, Read AP. Human Molecular Genetics. 2nd ed. New york: Wiley-Liss, 1999
4. 左伋. 医学遗传学. 第4版. 北京:人民卫生出版社,2004
5. Jorde LB. Medical Genetics. 2nd ed. St Louis: Mosby, 2000
6. Nussbaum RL. Thompson & Thompson Genetics in Medicine. 7th ed, Saunders, 2007

第五章 单基因遗传病

存在于生殖细胞或受精卵中的突变基因,按一定方式在上下代之间进行传递,其所携带的突变的遗传信息经过表达则可形成具有一定异常性状的疾病,这类疾病称为单基因遗传病(single-gene disorder, monogenic disorder)。单基因遗传病的发生主要受一对等位基因控制,其传递方式遵循孟德尔遗传规律。在单基因遗传病中,根据决定该疾病的基因所在染色体不同(常染色体或性染色体),以及该基因性质的不同(显性或隐性),可将人类单基因遗传病分为三种主要遗传方式:①常染色体遗传(autosomal inheritance),其中又包括常染色体显性遗传和常染色体隐性遗传;②X 连锁遗传(X-linked inheritance),包括 X 连锁显性遗传和 X 连锁隐性遗传;③Y 连锁遗传。线粒体基因组缺陷所引起的疾病虽然多为单基因遗传,但由于它属于细胞核外遗传,所以将在第九章中介绍。

研究人类疾病或性状的遗传规律不可能采用动植物性状遗传研究所普遍采用的杂交试验方法,因而需要建立一些研究人类遗传方式的特殊方法。系谱分析法(pedigree analysis)是其中最常见的方法。先证者(proband)是某个家族中第一个被医师或遗传研究者发现的罹患某种遗传病的患者或具有某种性状的成员。所谓系谱(pedigree)(或系谱图)是从先证者入手,追溯调查其所有家族成员(直系亲属和旁系亲属)的数目、亲属关系及某种遗传病(或性状)的分布等资料,并按一定格式将这些资料绘制而成的图解(常用的系谱绘制符号见图 5-1)。系谱中不仅要包括具有某种性状或患有某种疾病的个体,也应包括家族的正常成员。根据调查资料再绘制成系谱图,可以对这个家系进行回顾性分析,以便确定所发现的某一特定性状或疾病在这个家族中是否有遗传因素的作用及其可能的遗传方式,从而为其他具有相同遗传病的家系或患者的诊治提供依据。

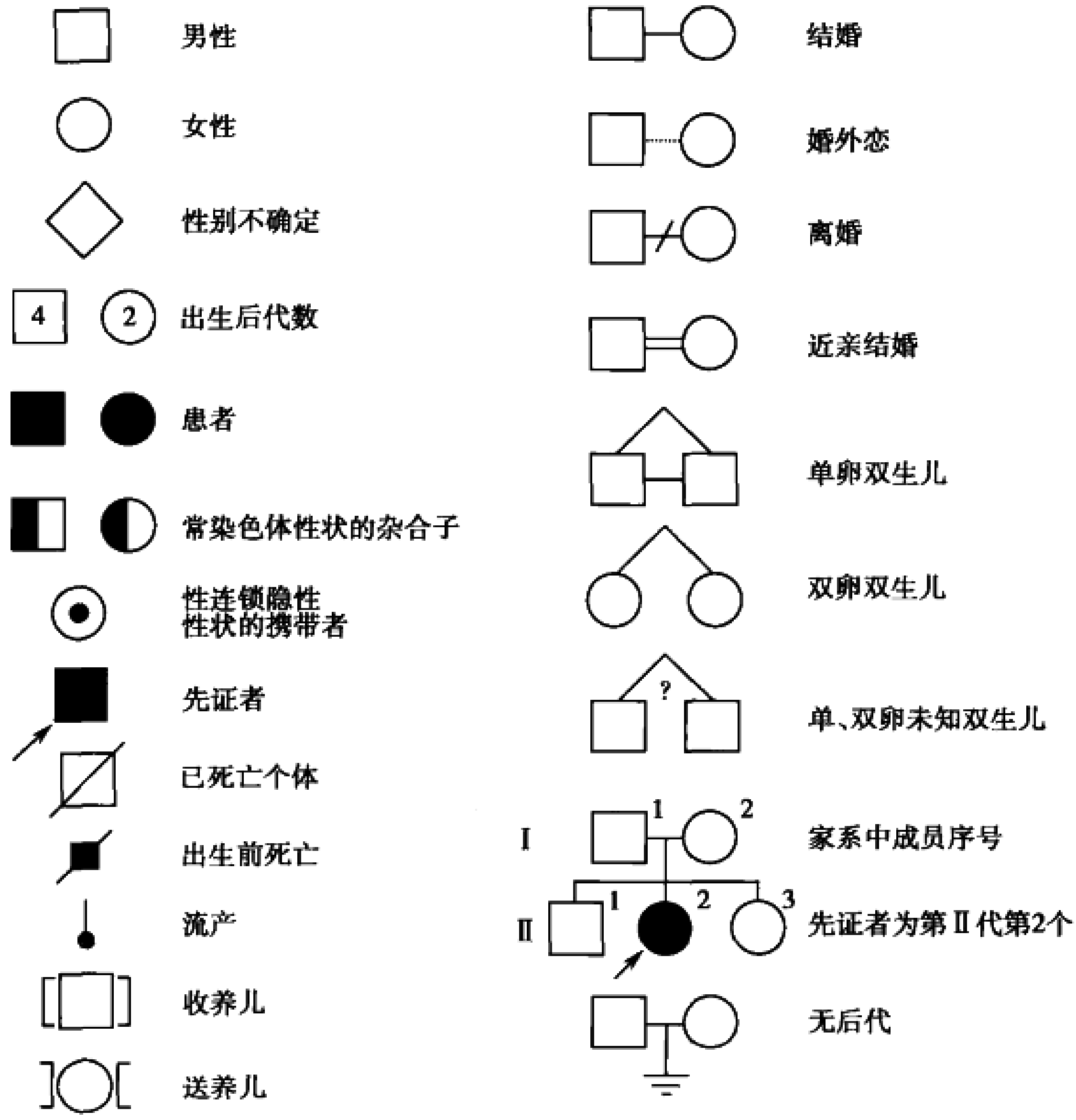


图 5-1 常用的系谱绘制符号

必须强调的是,在对某一种遗传性状或遗传病作系谱分析时,仅依据一个家族的系谱资料往往不能反映出该病的遗传方式及其特点。通常需要将多个具有相同遗传性状或遗传病的家族的系谱作综

合分析（统计学分析），才能比较准确而可靠地作出判断。

第一节　常染色体显性遗传病的遗传

人类的许多性状呈常染色体显性遗传（autosomal dominant inheritance，AD），例如，在决定人耳形态的三个主要性状中，长耳壳对短耳壳为显性，宽耳壳对狭耳壳为显性，有耳垂对无耳垂为显性。也就是说，长耳壳、宽耳壳、有耳垂等性状都受显性基因控制，短耳壳、狭耳壳、无耳垂等性状均为隐性基因所控制。而在人类疾病中也有不少疾病呈常染色体显性遗传。

一、常染色体显性遗传病举例：Huntington 舞蹈病

Huntington 舞蹈病（Huntington chorea；OMIM143100）又称遗传性舞蹈病（hereditary chorea），是一种典型的常染色体显性遗传病（图 5-2）。该病常于 30 ~ 45 岁时缓慢起病。患者有大脑基底神经节的变性，可引起广泛的脑萎缩，病变主要位于尾状核、豆状核（主要是壳核）和额叶。临床表现为进行性加重的舞蹈样不自主运动（不能控制的痉挛和书写动作）和智能障碍。患者舞蹈样运动的动作快，而且累及全身肌肉，但以面部和上肢最明显。每一阵舞蹈运动间有一较长间歇期，不自主运动在睡眠时消失。随着病情加重，可出现语言不清，甚至发音困难；精神症状常在不自主运动发生 1 ~2 年或数年后出现；智能障碍为进行性加重，最终出现痴呆。

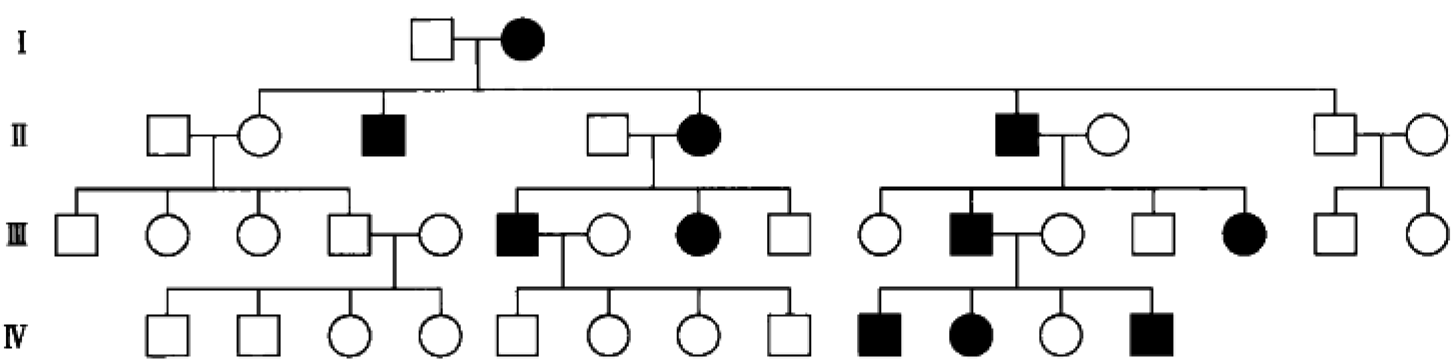


图 5-2　一个 Huntington 舞蹈病系谱

该病基因定位于 4p16.3，正常情况下，基因编码了一种称为 huntingtin 的蛋白，在疾病状态下，由于 *huntingtin* 基因 5′端（CAG）$_n$ 的动态突变，且（CAG）$_n$ 重复的多少与发病的早晚、疾病的严重程度成正比。正常人的（CAG）$_n$ 重复次数在 9 ~ 34 次，Huntington 舞蹈病患者大于 36 次，最多超过 120 次。发生了突变的 huntingtin 蛋白其羧基端串联重复的谷氨酰胺（polyQ）数量大大增加，使之在进入细胞核后不能正常地发挥调节基因转录的作用，而是相互聚集，形成核内包涵体（nuclear inclusion），最终导致神经元变性，死亡（图 5-3）。

其他常见且主要的常染色体显性遗传病见表 5-1。

二、婚配类型与子代发病风险

如果用 A 代表决定某种显性性状的基因，用 a 代表其相应的隐性等位基因，那么在完全显性（complete dominance）的情况下，杂合子 Aa 与显性纯合子 AA 的表型完全相同，即在杂合子 Aa 中，显性基因 A 的作用完全表现出来，而隐性基因 a 的作用被完全掩盖，从而使杂合子表现出与显性纯合体完全相同的性状。

短指（趾）症（brachydactyly A1；OMIM112500）是一种常染色体完全显性遗传的典型例子。它的主要症状是患者指骨（或趾骨）短小或缺失，致使手指（或足趾）变短（图 5-4）；致病基因定位于 2q33-q35。

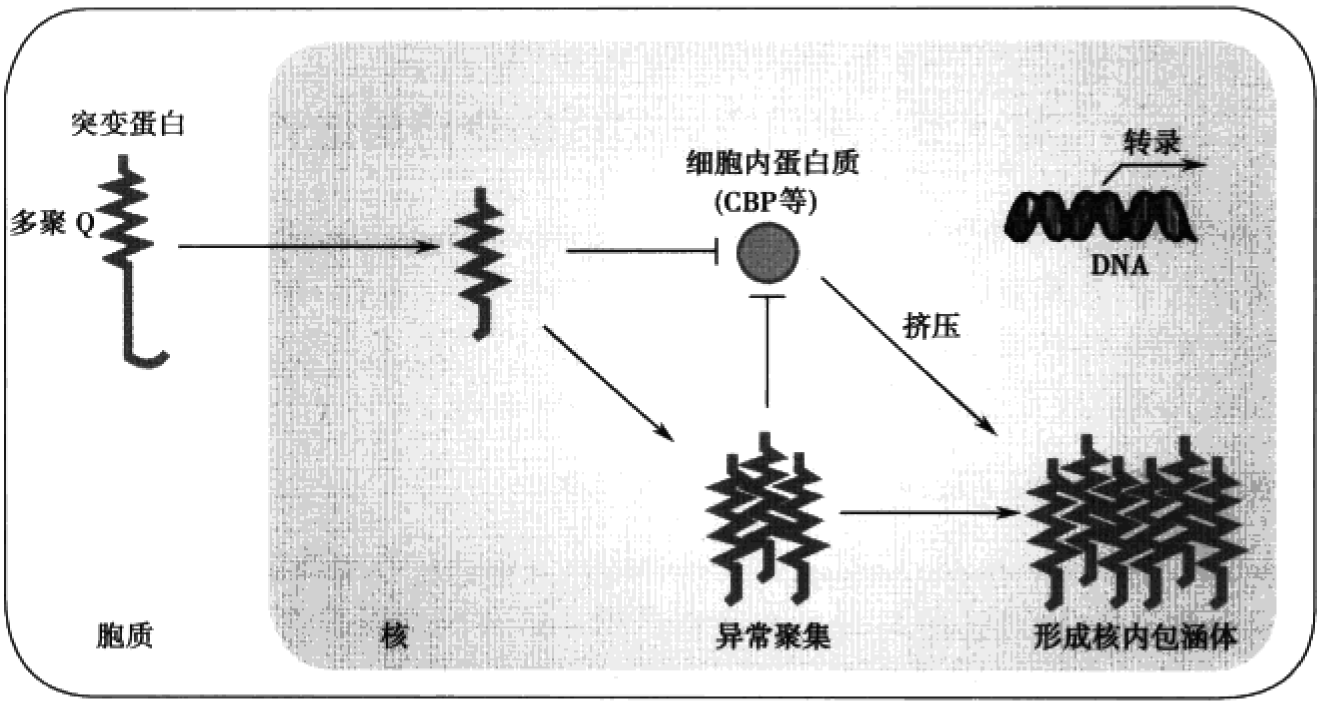


图 5-3 突变的 huntingtin 的致病过程

表 5-1 一些常染色体显性遗传病举例

疾病中文名称	疾病英文名称	OMIM	染色体定位
家族性高胆固醇血症	familial hypercholesterolemia	143890	19p13.2
遗传性出血性毛细血管扩张	hereditary hemorrhagic telangiectasia	187300	9q34.1
遗传性球形红细胞症	elliptocytosis	130500	1p36.2-p34
急性间歇性卟啉症	porphyria, acute intermittent	176000	11q23.3
迟发性成骨发育不全症 I 型	osteogenesis imperfecta, type I	166200	17q21.31-q22
成年多囊肾病	polycystic kidney disease, adult	173900	16p13.3-p13.12
α 地中海贫血	alpha-thalassemias	141800	16pter-p13.3
短指(趾)症 A1 型	brachydactyly, type A1	112500	2q35-q36
特发性肥大性主动脉瓣下狭窄	supravalvular aortic stenosis	185500	7q11.2
遗传性巨血小板病，肾炎和耳聋	Fechtner syndrome	153640	22q11.2
Noonan 综合征	Noonan syndrome	163950	12q24.1
神经纤维瘤 I 型	neurofibromatosis, type I	162200	17q11.2
结节性脑硬化	tuberous sclerosis	191100	16p13.3, 9q34
多发性家族性结肠息肉症	adenomatous polyposis of the colon	175100	5q21-q22
Peutz-Jeghers 综合征	Peutz-Jeghers syndrome	175200	19p13.3
Von Willebrand 病	Von Willebrand disease	193400	12p13.3
肌强直性营养不良	dystrophia myotonica	160900	19q13.2-q13.3

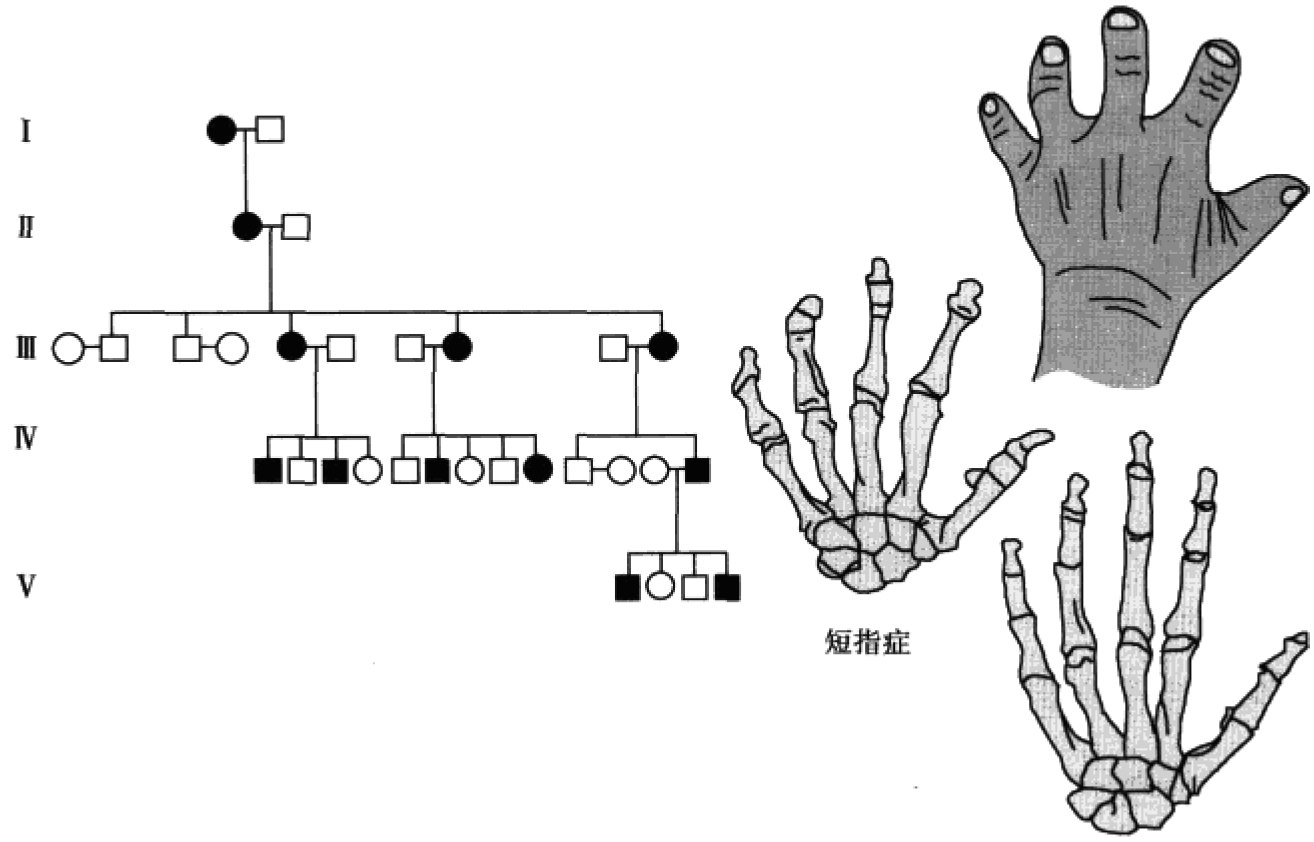


图 5-4 一个短指症家族的系谱

假设决定短指的基因为显性基因 A，正常指为隐性基因 a，则短指症患者基因型应为 AA 或 Aa。显性基因 A 在杂合状态下是完全显性的，因而在临床上，基因型为 AA 与 Aa 的患者在表型上完全不能区分。但实际上绝大多数短指症的基因型为 Aa，而不是 AA。因为根据分离律，基因型 AA 中的两个 A，必然一个来自父方，一个来自母方。这样，只有当父母都是短指症患者时，才能可能生出 AA 型子女，而这种婚配机会在实际生活中毕竟是很少的，并且显性致病基因在群体中的频率(p)很低，约为 1/1 000～1/100，根据遗传平衡定律(Hardy-Weinberg 定律，详见第七章)，显性纯合子短指症患者 AA 的频率(p^2)则更低，大约 1/1 000 000～1/10 000，而杂合子短指症患者 Aa 的频率(2pq)可达 1/500～1/50，故绝大多数短指症患者为 Aa。如果患者 Aa 与正常人 aa 婚配，其所生子女中，大约有 1/2 是患者(图 5-5)，也就是说，这对夫妇每生一个孩子，都有 1/2 的可能性生出短指症的患儿。

		亲代(Aa)(患者)	
		A	a
亲代(aa)(正常)	a	Aa	aa
	a	Aa	aa

子代表现型	正常(aa)		患者(Aa)
概率	1/2		1/2
概率比	1	:	1

图 5-5 常染色体显性遗传病杂合子患者与正常人婚配图解

三、常染色体完全显性遗传的特征

从以上典型的病例可见，常染色体完全显性遗传的典型遗传方式有如下特点：①由于致病基因位于常染色体上，因而致病基因的遗传与性别无关，即男女患病的机会均等；②患者的双亲中必有一个

为患者，但绝大多数为杂合子，患者的同胞中有1/2的可能性也为患者；③系谱中可见本病的连续传递，即通常连续几代都可以看到患者；④双亲无病时，子女一般不会患病（除非发生新的基因突变）。

根据这些特点，临床上可对常染色体完全显性的遗传病进行发病风险的估计。例如夫妇双方中有一人患病（杂合子），那么子女患病的可能性为1/2；两个患者（均为杂合子）婚配，则子女患病的可能性为3/4。

第二节　常染色体隐性遗传病的遗传

由于常染色体隐性遗传（autosomal recessive inheritance，AR）病的致病基因为隐性基因，所以只有隐性纯合子才会发病。在杂合子时，隐性致病基因的作用被其显性基因所掩盖，而不表现相应的疾病，表型与正常人相同，但是却可将致病基因遗传给后代。这种表型正常而带有致病基因的杂合子，称为携带者（carrier）。白化病、先天性聋哑、先天性肌弛缓等都属于此种遗传方式。

一、常染色体隐性遗传病举例：α_1-抗胰蛋白酶缺乏症

α_1-抗胰蛋白酶缺乏症（α_1-AT deficiency；OMIM 107400）是一种由 *α_1-AT* 基因突变引起的常染色体隐性遗传病，其特征是血清中 α_1-AT 水平下降。突变型最初在北欧、高加索人种中发现，以后传遍欧洲，又由于移民传至美国和其他国家。

最常见的 *α_1-AT* 基因突变型是S型和Z型，都属于单碱基改变型。S型较Z型更常见，还有一种无效型（null-null）很少见，其他突变型更罕见。S突变型是 *α_1-AT* 基因的外显子Ⅲ中发生单个碱基取代，致使合成的 α_1-AT 分子中的264Glu被264Val代替。这使得 α_1-AT 分子中的离子键丢失，改变了 α_1-AT 分子内部的结构，分子稳定性受到影响。Z突变型是 *α_1-AT* 外显子E中发生单个碱基取代，其合成的 α_1-AT 分子中的342Glu被342Lys代替，这也使离子键丢失，α_1-AT 分子的稳定性也受影响。无效突变个体的 α_1-AT 合成细胞中，α_1-AT mRNA 转录物缺失，表型的血清中完全测不到 α_1-AT，Z型和无效型个体都易发生肺气肿。

其他一些常见且主要的常染色体隐性遗传病见表5-2。

表5-2　常染色体隐性遗传病举例

疾病中文名称	疾病英文名称	OMIM	染色体定位
镰状细胞贫血	sickle cell anemia	603 903	11p15.5
家族性黑矇性白痴	Tay-Sachs disease	272 800	15q23-q24
β地中海贫血	beta-thalassemias	141 900	11p15.5
同型胱氨酸尿症	homocystinuria	236 200	21q22.3
苯丙酮尿症	phenylketonuria	261 600	12q24.1
丙酮酸激酶缺乏症	pyruvate kinase deficiency of erythrocyte	266 200	1q21
尿黑酸尿症	alkaptonuria	203 500	3q21-q23
Friedreich 家族性共济失调	Friedreich ataxia	208 900	11q22.3
Bardet-Biedl 综合征	Bardet-Biedl syndrome	209 900	20p12
半乳糖血症	galactosemia	230 400	9p13
肝豆状核变性	Wilson disease	277 900	13q14.3-q21.1
黏多糖累积症Ⅰ型	mucopolysaccharidosis type Ⅰ	252 800	4p16.3
先天性肾上腺皮质增生	adrenal hyperplasia，congenital	201 910	6p21.3
血浆活酶前体缺乏症	PTA deficiency	264 900	4q35
囊性纤维变性	cystic fibrosis	219 700	7q31.2
血色素沉着症	hemochromatosis	235 200	6p21.3

二、婚配类型及子代发病风险

在常染色体隐性遗传病家系中最常见的是两个杂合子(Aa×Aa)的婚配，每胎孩子得病的概率是0.25，在患者的表现型正常同胞中杂合子占2/3，因此该类婚配家庭的子女中将有1/4得病(图5-6)。

杂合亲代(Aa)(携带者) \ 杂合亲代(Aa)(携带者)	A	a
A	AA	Aa
a	Aa	aa

子代表现型	正常(AA)	表型正常的携带者(Aa)	患者(aa)
概率	1/4	2/4	1/4
概率比	1 :	2 :	1

图5-6 常染色体隐性遗传病杂合子相互婚配图解

实际上，人群中最多的婚配类型应该是杂合子与正常人婚配(Aa×AA)，子代表现型全部正常，但其中将有1/2是携带者(图5-7)。

纯合亲代(AA)(正常) \ 杂合亲代(Aa)(携带者)	A	a
A	AA	Aa
A	AA	Aa

子代表现型	正常(AA)	表型正常的携带者(Aa)
概率	1/2	1/2
概率比	1 :	1

图5-7 常染色体隐性遗传病杂合子与正常人婚配图解

杂合子与患者婚配(Aa×aa)可能发生于近亲婚配时，子代中将有一半为患者，另一半为携带者(图5-8)。这种家系由于连续两代出现患者，子代比例模拟显性遗传格局，称为类显性遗传(quasidominant inheritance)，不易与常染色体显性遗传区分。在近亲婚配家庭中出现这种遗传格局时，应考虑常染色体隐性遗传的可能性。

患者相互婚配(aa×aa)时，子女无疑将全部受累。由于隐性致病基因少见，这种婚配的可能性极少，只有在发病率高的隐性遗传病中才能见到。

三、常染色体隐性遗传的遗传特征

一般认为，常染色体隐性遗传的典型系谱(图5-9)有如下特点：①由于基因位于常染色体上，所以它的发生与性别无关，男女发病机会相等；②系谱中患者的分布往往是散发的，通常看不到连续传递现象，有时在整个系谱中甚至只有先证者一个患者；③患者的双亲表型往往正常，但都是致病基因的携带者，此时出生患儿的可能性约占1/4，患儿的正常同胞中有2/3的可能性为携带者；④近亲婚配时，子女中隐性遗传病的发病率要比非近亲婚配者高得多。这是由于他们来自共同的祖先，往往具

		杂合亲代(Aa)(携带者)	
		A	a
纯合亲代(aa)(患者)	a	Aa	aa
	a	Aa	aa

子代表现型	患者(aa)		表型正常的携带者(Aa)
概率	1/2		1/2
概率比	1	:	1

图 5-8 常染色体隐性遗传病患者与杂合子婚配图解

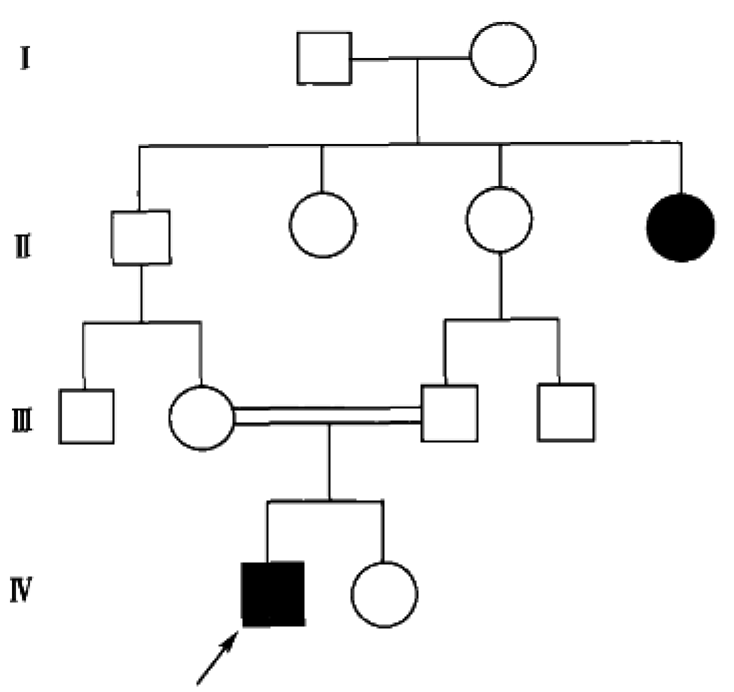


图 5-9 常染色体隐性遗传的典型系谱

有某种共同的基因。

四、常染色体隐性遗传病分析时应注意的问题

在临床上所看到的常染色体隐性遗传病家系中,可出现患者人数占其同胞人数的比例高于理论上的 1/4 的现象,这是由于存在有选择偏倚(selection deviation)。因为,在常染色体隐性遗传病家系中,父母一方患病时,子女中有 1 个或 1 个以上患病的人或无患病的人均被确认,所得数据较为完整,称为完全确认(complete ascertainment),这个数据将接近于 1/1。当一对夫妇都是某种常染色体隐性遗传病基因的携带者,他们一生又只生一个或少数几个孩子,如果他们所生的孩子没有病(概率为 3/4),他们就不会来就诊,也就不会列入医师的统计范畴,造成无患病子女家系的漏检,称为不完全确认或截短确认(truncate ascertainment);因此如果患了某种常染色体隐性遗传病 (概率为 1/4) 并到医院就诊, 显然其比例就偏高, 对于在只生一个或第 1 个孩子即为患者的家庭中, 医师所统计到的比例为 100%。事实上, 在生育子女数目更多的家庭中, 也存在着这种选择偏倚。因此在计算常染色体隐性遗传病患者同胞的发病比例时, 必须采用一种校正的方法。

目前常用的方法是 Weinberg 先证者法,其校正公式为 $C=\frac{\sum a(r-1)}{\sum a(s-1)}$。这里 C 为校正比例;a 为先证人数;r 为同胞中的受累人数;s 为同胞人数。

例如,对 11 个苯丙酮尿症患者家庭的调查中发现,有 4 个先证者仅有同胞 1 人,并且同为患者;有 3 个先证者有同胞 2 人,其中 2 家各仅有 1 名患者,另 1 家有 2 名患者;有 3 个先证者有同胞 3 人,其中 2 家各有 1 名患者,另 1 家有 2 名患者;有 1 个先证者有同胞 4 人,其中有 2 名患者。在总共 23 名同胞中,发病者 14 人,发病比例为 14/23 = 0. 6087,大大高于期望值 1/4(0. 25)。

如按校正公式 $C=\frac{\sum a(r-1)}{\sum a(s-1)}$进行校正,则先列表 5-3 如下,再将表中的数值代入公式:C = 3/12 = 0. 25,校正的数据表明,观察到的苯丙酮尿症患者同胞中的发病比例完全符合常染色体隐性遗传病的发病比例,即 1/4。

表5-3 苯丙酮尿症 Weinberg 先证者法校正表

s	r	a	a(r-1)	a(s-1)
1	1	1	0	0
1	1	1	0	0
1	1	1	0	0
1	1	1	0	0
2	1	1	0	1
2	1	1	0	1
2	2	1	1	1
3	1	1	0	2
3	1	1	0	2
3	2	1	1	2
4	2	1	1	3
Σ23	14	11	3	12

第三节 X 连锁显性遗传病的遗传

由性染色体的基因所决定的性状在群体分布上存在着明显的性别差异是性连锁遗传的特征。如果决定某种性状或疾病的基因位于 X 染色体上，并且此基因对其相应的等位基因来说是显性的，这种遗传病的遗传方式称之为 X 连锁显性遗传（X-linked dominant inheritance，XD）。

男性只有一条 X 染色体，其 X 染色体上的基因在 Y 染色体上缺少与之对应的等位基因，因此男性只有成对基因中的一个成员，故称半合子（hemizygote），其 X 染色体上有此基因才表现出相应性状或疾病。而女性有两条 X 染色体，其中任何一条 X 染色体上有此基因，都可以表现出相应的性状。因此 X 连锁显性遗传病的发病率女性要比男性约高一倍，但病情男重于女。

一、X 连锁显性遗传病举例：抗维生素 D 佝偻病

抗维生素 D 佝偻病（vitamin D-resistant rickets；OMIM307800）又称低磷酸盐血症（hypophosphatemia），是一种以低磷酸盐血症导致骨发育障碍为特征的遗传性骨病（图 5-10）。患者由于肾小管对磷酸盐再吸收障碍，从而血磷下降，尿磷增多，肠道对磷、钙的吸收不良而影响骨质钙化，形成佝偻

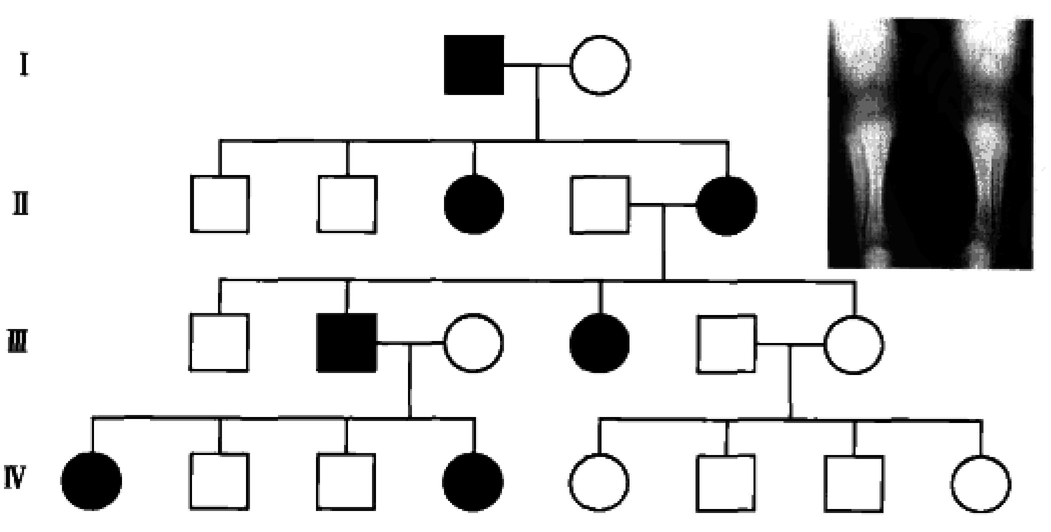


图 5-10 一抗维生素 D 佝偻病系谱

病。患儿多于1周岁左右发病，最先出现的症状为O形腿，严重的有进行性骨骼发育畸形、多发性骨折、骨疼、不能行走、生长发育缓慢等症状。从临床观察，女性患者的病情较男性患者轻，少数只有低磷酸盐血症，而无佝偻病的骨骼变化，这可能是女性患者多为杂合子，其中正常X染色体的基因还发挥一定的作用。

该病基因已定位于Xp22.1-p22.2，基因已被克隆，称为*PHEX*(phosphate regulated gene with homologies to endopeptidases, on the X chromosome)，该基因编码749氨基酸残基。缺失和单个碱基置换是导致疾病发生的主要原因。

其他一些常见和主要的X连锁显性遗传病见表5-4。

表5-4 X连锁显性遗传病举例

疾病中文名称	疾病英文名称	OMIM	染色体定位
口面指综合征Ⅰ型	Orofaciodigital syndrome, type Ⅰ	311200	Xp22.3-p22.2
高氨血症Ⅰ型(鸟氨酸氨甲酰基转移酶缺乏)	ornithine transcarbamylase deficiency	311250	Xp21.1
Alport综合征	Alport syndrome	301050	Xq22.3
色素失调症	incontinentia pigmenti	308300	Xq28

二、婚配类型和子代发病风险

X连锁显性遗传病的致病显性突变基因在X染色体上，只要一条X染色体上有此突变基因(即女性杂合子或男性半合子)即可致病。如果调查的只是女性患者的子代，这时的系谱格局不能与常染色体显性遗传相区别，关键在于没有父到子传递；X连锁显性遗传病患者女性多于男性，约呈2:1(图5-11)。

杂合女性患者(X^RX) \ 正常男性(XY)	X	Y
X^R	X^RX	X^RY
X	XX	XY

子代表现型	正常女性(XX)	杂合女性患者(X^RX)	正常男性(XY)	半合子男性患者(X^RY)
概率	1/4	1/4	1/4	1/4
概率比	1 :	1 :	1 :	1

图5-11 X连锁显性遗传病杂合女性患者与正常男性婚配图解

半合子男性患者(X^RY)与正常女性(XX)婚配。由于男性患者的X^R一定给女儿，因此女儿都将是患者，儿子全部为正常(图5-12)。

三、X连锁显性遗传的遗传特征

X连锁显性遗传的典型系谱遗传方式有如下特点：①人群中女性患者比男性患者约多一倍，前者病情常较轻；②患者的双亲中必有一名是该病患者；③男性患者的女儿全部都为患者，儿子全部正常；④女性患者(杂合子)的子女中各有50%的可能性是该病的患者；⑤系谱中常可看到连续传递现象，这点与常染色体显性遗传一致。

半合子男性患者(X^RY)

正常女性(XX)	X^R	Y
X	X^RX	XY
X	X^RX	XY

子代表现型	杂合女性患者(X^RX)		正常男性(XY)
概率	1/2		1/2
概率比	1	:	1

图 5-12　X 连锁显性遗传病半合子男性患者与正常女性婚配图解

第四节　X 连锁隐性遗传病的遗传

如果决定某种性状或疾病的基因位于 X 染色体上，且为隐性基因，这种基因的遗传方式称为 X 连锁隐性遗传(X-recessive inheritance，XR)，以 XR 方式遗传的疾病称为 X 连锁隐性遗传病。

一、X 连锁隐性遗传病举例：Duchenne 型肌营养不良症

Duchenne 型肌营养不良症(Duchenne muscular dystrophy，DMD OMIM310200)又称假性肥大性肌营养不良。本症是一种非神经性病因所造成的疾病，而是由肌肉细胞本身随着时间及年龄渐进性损伤与萎缩的疾病。这群肌肉疾病大部分是基因有特定的突变所致，因此在临床上，具有特殊的遗传性质。本病发病率约为 1/3500，由于它是 X 连锁隐性遗传，因此只有男孩罹患此症，而女孩通常不发病但携带有隐性突变的基因。患者多于 4～5 岁发病，初期感觉走路笨拙，易于跌倒，不能奔跑及登楼，站立时脊椎前凸，腹部挺出，两足撇开，步行缓慢摇摆，呈特殊的"鸭步"步态，当仰卧起立时非常困难，必先翻身俯卧，再双手攀缘两膝，逐渐向上支撑起立(Gower 征)(图 5-13)。后期患者双侧腓肠肌假性肥大，病变肌纤维肿胀，粗细不等，散布于正常纤维之间，肌核增大增多且排列成链。残存的肌纤维间有结缔组织增生及脂肪沉淀。

家系调查和细胞遗传学研究证明，*DMD* 基因是由于定位于 Xp21.2 的抗肌萎缩蛋白(dystrophin)

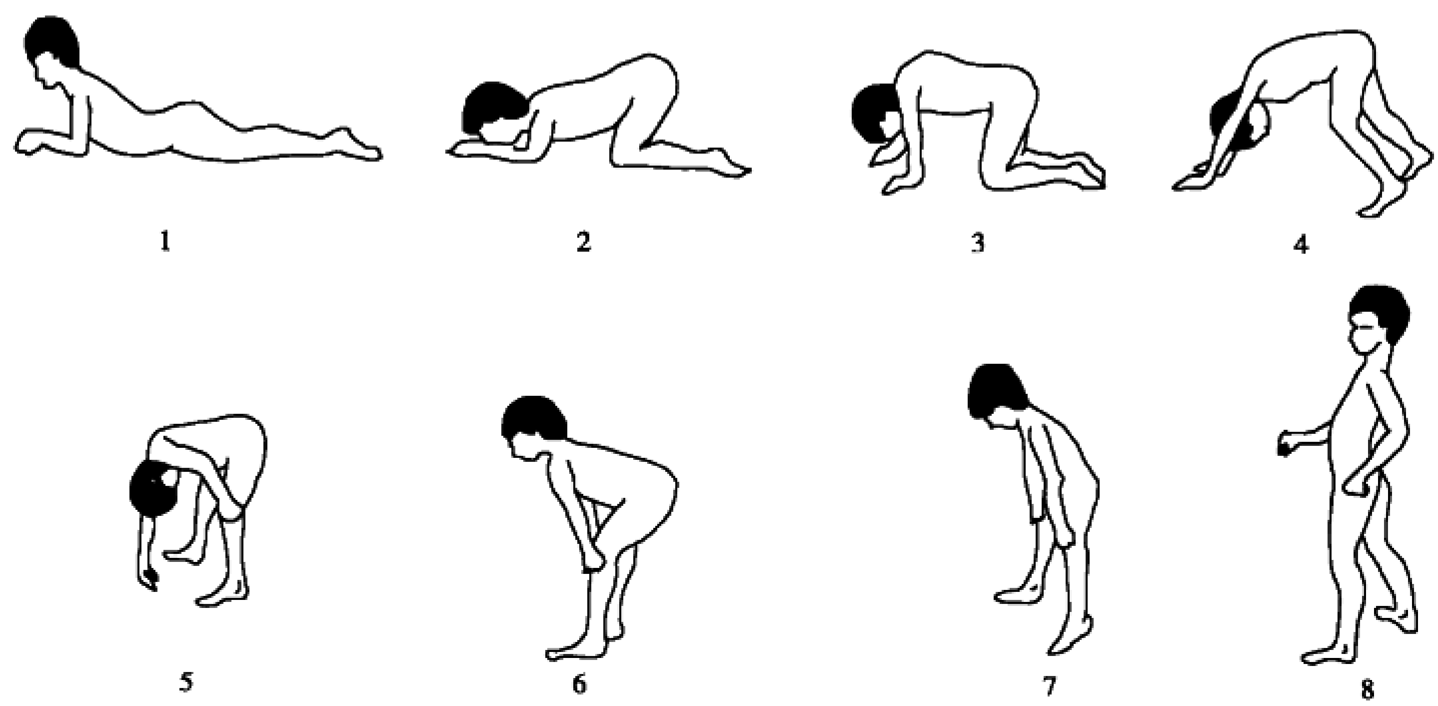


图 5-13　Duchenne 肌营养不良患儿从俯卧位靠双手支撑爬起来的动作(Gowers 征)

基因缺陷，使其产物 dystrophin 蛋白不能在肌细胞膜上正常表达所致。因此主要表现为肌肉变性、萎缩及进行性肌无力等。Dystrophin 基因长达 2400kb，79 个外显子，cDNA 全长 14kb，编码的肽链含 3685 残基，分子量为 427 000。是目前发现的人类最大的基因。迄今人们已发现 Dystrophin 基因突变的类型有多种，其中缺失占 60%、重复及点突变占 10%、点突变及微小缺失或常规方法检测不出的微小重复约占 30%。

其他一些常见且主要的 X 连锁隐性遗传病见表 5-5。

表 5-5 X 连锁隐性遗传病举例

疾病中文名称	疾病英文名称	OMIM	染色体定位
色盲	colorblindness	303 800	Xq28
睾丸女性化	androgen insensitivity syndrome	300 068	Xq11-q12
鱼鳞癣	ichthyosis	308 100	Xp22. 32
Lesch-Nyhan 综合征	Lesch-Nyhan syndrome	300 322	Xq26-q27. 2
眼白化病 I 型	albinism, ocular, type I	300 500	Xp22. 3
Hunter 综合征	mucopolysaccharidosis type II	309 900	Xq28
无丙种球蛋白血症 I 型	immunodeficiency with hyper-IgM, type I	308 230	Xq26
Fabry 病(糖鞘脂贮积症)	Fabry disease	301 500	Xq22
Wiskott-Aldrich 综合征	Wiskott-Aldrich syndrome	301 000	Xp11. 23-p11. 22
G-6-PD 缺乏症	glucose-6-phosphate dehydrogenase	305 900	Xq28
肾性尿崩症	diabetes insipidus, nephrogenic, X-linked	304 800	Xq28
慢性肉芽肿病	granulomatous disease	306 400	Xp21. 1
血友病 B	hemophilia B	306 900	Xq27. 1-q27. 2
无汗性外胚层发育不良症	ectodermal dysplasia	305 100	Xq12-q13. 1

二、婚配类型和子代发病风险

在 X 连锁隐性遗传家系中最常见的是表现型正常的杂合子携带者女性(X^dX)与正常男性婚配，子代中将有半数儿子受累，半数女儿为携带者(图 5-14)。

		正常男性(XY)	
		X	Y
携带者女性(X^dX)	X^d	X^dX	X^dY
	X	XX	XY

子代表现型	正常女性(XX)	携带者女性(X^dX)	正常男性(XY)	半合子男性患者(X^dY)
概率	1/4	1/4	1/4	1/4
概率比	1 :	1 :	1 :	1

图 5-14 X 连锁隐性遗传病杂合子女性与正常男性婚配图解

半合子男性患者(X^dY)与正常女性婚配，所有儿子和女儿的表现型都正常，但父亲的 X^d 一定给了女儿，因此所有女儿均为杂合子携带者(图 5-15)。

半合子男性患者(X^dY)

		X^d	Y
正常女性(XX)	X	X^dX	XY
	X	X^dX	XY

子代表现型	携带者女性(X^dX)		正常男性(XY)
概率	2/4		2/4
概率比	1	:	1

图 5-15 X 连锁隐性遗传病半合子男性与正常女性婚配图解

三、X 连锁隐性遗传的遗传特征

X 连锁隐性遗传的典型遗传方式有如下特点:①人群中男性患者远较女性患者多,系谱中往往只有男性患者;②双亲无病时,儿子可能发病,女儿则不会发病;儿子如果发病,母亲肯定是一个携带者,女儿也有 1/2 的可能性为携带者;③男性患者的兄弟、外祖父、舅父、姨表兄弟、外甥、外孙等也有可能是患者;④如果女性是一患者,其父亲一定也是患者,母亲一定是携带者。

第五节 Y 连锁遗传病的遗传

如果决定某种性状或疾病的基因位于 Y 染色体,那么这种性状(基因)的传递方式称为 Y 连锁遗传(Y-linked inheritance)。

Y 连锁遗传的传递规律比较简单,具有 Y 连锁基因者均为男性,这些基因将随 Y 染色体进行传递,父传子、子传孙,因此称为全男性遗传。

目前已经知道的 Y 连锁遗传的性状或遗传病比较少,肯定的有 H-Y 抗原基因、外耳道多毛基因和睾丸决定因子基因等。图 5-16 为一个外耳道多毛症系谱。该系谱中全部男性均有此性状,即到了青春期,外耳道中可长出 2~3cm 的成丛黑色硬毛,常可伸出到耳孔之外。系谱中所有女性均无此症状。

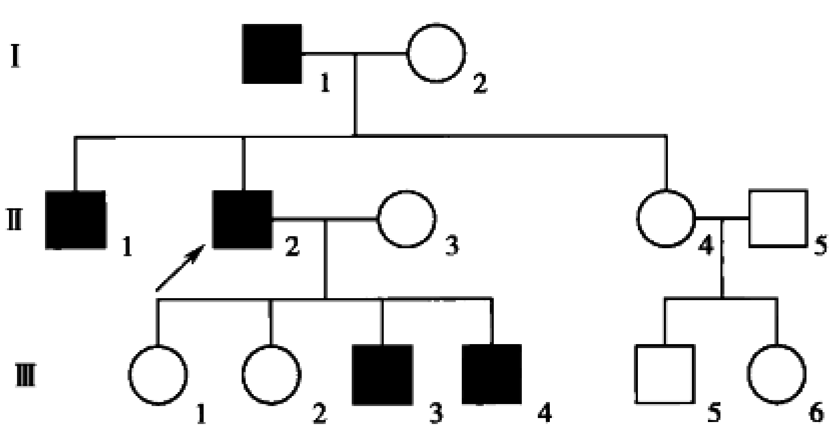


图 5-16 外耳道多毛症的系谱

以上介绍了单基因遗传的几种主要遗传方式及特点。对于某一遗传性状或某种疾病来说,通过多个家系的调查和系谱分析,有助于对该性状或疾病的遗传方式做出初步的估计和预测子女的发病风险。

第六节 影响单基因遗传病分析的因素

根据基因突变的性质,通常把与其所控制的相应表型分为显性遗传和隐性遗传两大类。理论上,两者在群体中呈现出各自的分布规律,但某些突变基因性状的遗传存在着许多例外情况。

一、表现度

表现度(expressivity)也称为表现变异性(variable expressivity),是基因在个体中的表现程度,或者说具有同一基因型的不同个体或同一个体的不同部位,由于各自遗传背景的不同,所表现的程度可有显著的差异。例如常染色体显性遗传的成骨发育不全症,它以耳聋、蓝色巩膜、骨质脆弱以至易于骨折为主要症状。由于表现度的不同,有的只表现为蓝色巩膜;有的除蓝色巩膜外,还表现为耳聋;严重者除三大症状全部表现外还有牙齿半透明、指甲发育不全等症状。

多指(趾)症致病基因可以表现为指数多少的不一;桡侧多指与尺侧多指不一;手多指与脚多趾的不一;或软组织的增加和掌骨的增加程度不一等。而这些差异既可出现在不同个体,也可出现在同一个体的不同部位。

二、外显率

外显率(penetrance)是某一显性基因(在杂合状态下)或纯合隐性基因在一个群体中得以表现的百分比。仍以多指(趾)症为例,在调查某一群体后,推测具有该致病基因的个体数为 25 人,而实际具有多指(趾)表型的人数为 20 人。因此,所调查群体中该致病基因的外显率为 $20/25 \times 100\% = 80\%$。外显率等于 100% 时为完全外显(complete penetrance);低于 100% 时则为外显不全或不完全外显(incomplete penetrance)。当然某一基因的外显率不是绝对不变的,相反,它随着观察者所定观察标准的不同而变化。上述的多指(趾)症致病基因的外显率是以肉眼观察指(趾)的异常与否为标准的。若辅以 X 线摄影,就可发现因肉眼未看出而被认为不外显的"正常人"也有骨骼的异常。若以此为标准,则多指(趾)症致病基因的外显率将有所提高。

外显率与表现度是两个不同的概念,切不可混淆。其根本的区别在于前者阐明了基因表达与否,是个"质"的问题;而后者要说明的是在表达前提下的表现程度如何,是个"量"的问题。

三、拟表型

由于环境因素的作用使个体的表型恰好与某一特定基因所产生的表型相同或相似,这种由环境因素引起的表型称为拟表型(phenocopy),或表现型模拟。例如常染色体隐性遗传的先天性聋哑,与由于使用药物(链霉素)引起的聋哑,都有一个相同的表型,即聋哑。这种由于药物引起的聋哑即为拟表型。显然,拟表型是由于环境因素的影响,并非生殖细胞中基因本身的改变所致,因此,这种聋哑并不遗传给后代。

四、基因的多效性

基因的多效性(pleiotropy)是一个基因可以决定或影响多个性状。在生物个体的发育过程中,很多生理生化过程都是互相联系、互相依赖的。基因的作用是通过控制新陈代谢的一系列生化反应而影响到个体发育的方式,从而决定性状的形成。这些生化反应按照特定的步骤进行,每一基因控制一个生化反应。因此,一个基因的改变直接影响其他生化过程的正常进行,从而引起其他性状的相应改变。这方面的例子是很多的。如半乳糖血症是一种糖代谢异常症,患者既有智能发育不全等神经系统异常,还具有黄疸、腹水、肝硬化等消化系统症状,甚至还可出现白内障。造成这种多效性的原因,并不是基因真正地具有多重效应,而是基因产物在机体内复杂代谢的结果。下面可从两个方面进行分析,一是基因产物(蛋白质或酶)直接或间接控制和影响了不同组织和器官的代谢功能,即所谓的初级效应。上述的半乳糖血症即属此例;二是在基因初级效应的基础上通过连锁反应引起的一系列次级效应。例如镰状细胞贫血,由于存在异常血红蛋白(HbS)引起红细胞镰变,进而使血液黏滞度增加、局部血流停滞、各组织器官的血管梗死、组织坏死等,导致各种临床表现。这些临床表现都是初级效应(镰变)后的次级效应,这是基因多效性的另一原因。

五、遗传异质性

与基因多效性相反，遗传异质性（genetic heterogeneity）是同一性状可以由多个不同的基因或同一基因多种不同突变控制，前者称为基因座异质性（locus heterogeneity），后者称为等位基因异质性（allelic heterogeneity）。例如智能发育不全这种异常性状，可由半乳糖血症的基因控制，也可由苯丙酮尿症的基因、黑朦性白痴基因所决定。随着人类知识水平的不断提高，实验技术、分析手段愈加精细，就会在越来越多的病例中观察到遗传异质性。如临床上表现相似的糖原累积症，现在已发现多种类型，每种类型都有其自己的基因缺陷。

六、表型异质性

同一基因的不同突变有时会产生截然不同的表型，称为表型异质性（phenotypic heterogeneity）。例如编码酪氨酸激酶的 *RET* 基因若发生某种突变而失去功能，引起呈显性遗传的肠神经节发育缺陷，患者丧失肠动力导致严重慢性便秘（先天性巨结肠症）；*RET* 基因的另一种突变则引起激酶活性增高，引起呈显性遗传的甲状腺癌和肾上腺癌（多发性内分泌瘤 2A 型和 2B 型）；*RET* 基因的第三类突变则导致同一个体同时患有先天性巨结肠症和多发性内分泌瘤。

七、遗传早现

遗传早现（anticipation）是指一些遗传病（通常为显性遗传病）在连续几代的遗传中，发病年龄提前而且病情严重程度增加。例如，遗传性小脑性运动共济失调（Marie 型）综合征是一种常染色体显性遗传病，其发病年龄一般为 35～40 岁，临床表现早期为行走困难，站立时摇摆不定，语言不清；晚期下肢瘫痪。由图 5-17 可见 $Ⅰ_1$ 39 岁开始发病，$Ⅱ_2$ 38 岁开始发病，$Ⅲ_3$ 30 岁发病，而 $Ⅳ_1$ 23 岁就已瘫痪。在许多家系分析中，都可以发现这种遗传早现。

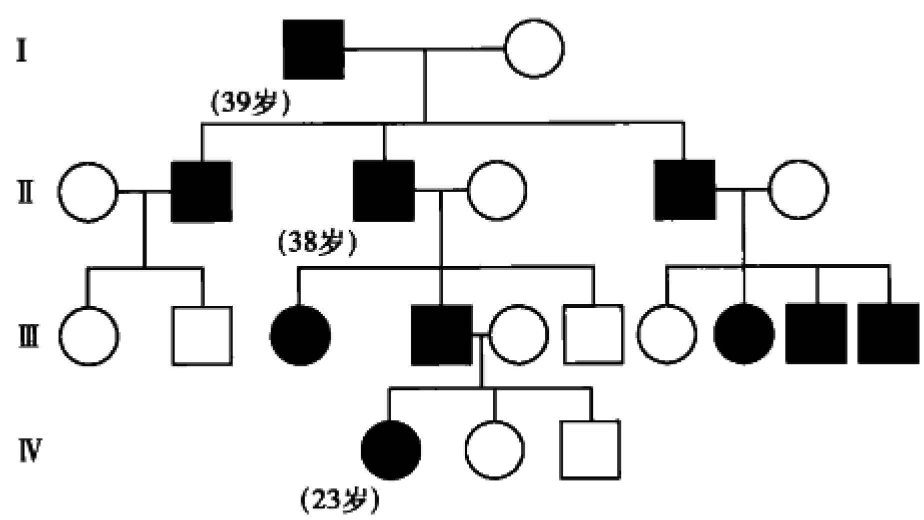


图 5-17　遗传性小脑运动共济失调（Marie 型）的系谱

八、不稳定性重复扩增

某些单基因遗传性状的异常改变或疾病的发生，是由于 DNA 分子中某些短串联重复序列发生不稳定性重复扩增（unstable repeat expansion），尤其是基因编码序列或侧翼序列的三核苷酸重复扩增所引起。因为这种三核苷酸的重复次数可随着世代交替的传递而呈现逐代递增的累加突变效应，故也被称为动态突变（dynamic mutation）。把由动态突变所引起的疾病，统称为三核苷酸重复扩增病（trinucleotide repeat expansion diseases，TREDs）。上述的遗传早现就是建立在这一特殊的遗传机制基础上的。

例如，在表现为 X 连锁隐性遗传特征的脆性 X 染色体综合征患者中，其 X 染色体 q27.3 处存

在有不稳定的易断裂脆性部位。利用限制性内切酶 Pst Ⅰ进行 X 染色体切割，可得到包括该脆性部位在内的限制性酶切片段。经序列分析表明，患者的这一限制性酶切片段中存在的$(CGG)_n$重复拷贝数可达 52～200 个；而在正常人则仅为 6～50 个。但$(CGG)_n$两边的侧翼序列却与正常人几无差异。

九、从性遗传

从性遗传（sex-conditioned inheritance）是位于常染色体上的基因，由于性别的差异而显示出男女性分布比例上的差异或基因表达程度上的差异。例如秃顶，是常染色体显性遗传，是一种从头顶中心向周围扩展的进行性对称性脱发。一般 35 岁左右开始出现秃顶，而且男性秃顶明显多于女性。这是因为杂合子男性表现秃顶；杂合子女性则不会表现。经研究表明秃顶基因能否表达还要受到雄性激素的影响。如果带有秃顶基因的女性，体内雄性激素水平升高也可出现秃顶。这点可作为诊断女性是否患某种疾病的辅助指标。因为肾上腺肿瘤可产生过量雄性激素，导致秃顶基因的表达。

再如原发性血色素病是一种由于铁质在体内器官的广泛沉积而引起损害的常染色体显性遗传病。男性的发病率远高于女性。究其原因，认为可能是由于女性月经、流产或妊娠等生理或病理性失血导致铁质丢失，减轻了铁质的沉积，故不易表现出症状。

十、限性遗传

限性遗传（sex-limited inheritance）是常染色体上的基因，由于基因表达的性别限制，只在一种性别表现，而在另一种性别则完全不能表现。这主要是由于解剖学结构上的性别差异造成的，也可能受性激素分泌方面的差异限制。如女性的子宫阴道积水症，男性的前列腺癌等。

十一、遗传印迹

越来越多的研究显示一个个体的同源染色体（或相应的一对等位基因）因分别来自其父方或母方，而表现出功能上的差异，因此当它们其一发生改变时，所形成的表型也有不同，这种现象称为遗传印迹（genetic imprinting）或基因组印迹（genomic imprinting）、亲代印迹（parental imprinting）。

在人类，由于印迹效应，一些单基因遗传病的表现度和外显率也受到突变基因亲代来源的影响。例如 Huntington 舞蹈病的基因如果经母亲传递，则其子女的发病年龄与母亲的发病年龄一样；如果经父亲传递，则其子女的发病年龄比父亲的发病年龄有所提前，在一些家系中，子女的发病年龄可能提前到 20 岁左右。但是这种发病年龄提前的父源效应经过一代传递即消失，早发型男性的后代仍然为早发型，而早发型女性的后代的发病年龄并不提前。其他疾病如脊髓小脑性共济失调、强直性肌萎缩和多发性神经纤维瘤等也存在有相似的印迹效应。

十二、延迟显性

杂合子在生命的早期，因致病基因并不表达或虽表达但尚不足以引起明显的临床表现，只在达到一定的年龄后才表现出疾病，这一显性形式称为延迟显性（delayed dominance）。例如 Huntington 舞蹈病常于 30～40 岁间发病，属于延迟显性的一个例子。

十三、X 染色体失活

Lyon 假说认为女性两条 X 染色体在胚胎发育早期就随机失活了其中的一条，即为 X 染色体失活（X inactivation，lyonization），因此女性的两条 X 染色体存在嵌合现象。平均说来，女性一半细胞表现父源染色体，另一半细胞表现母源染色体。如有一妇女为 X 连锁杂合子，预期半数细胞中带有突变

基因的那条 X 染色体失活，细胞是正常的，另外半数细胞中带有正常基因的那条 X 染色体失活，细胞将为突变型。但曾有报道，偶见 X 连锁隐性遗传的血友病或 Duchenne 型肌营养不良症的男性患者的杂合子母亲也可受累，这种 X 连锁隐性遗传的女性杂合子表现出临床症状是一种所谓"显示杂合子"(manifesting heterozygote)。这是因为女性 X 染色体有随机失活现象，机遇使她大部分细胞中带有正常基因的 X 染色体失活，而带有隐性致病基因的那条 X 染色体恰好有活性，从而表现出或轻或重的临床症状。

十四、不完全显性遗传

不完全显性(incomplete dominance)也称为半显性(semi-dominance)遗传。它是杂合子 Dd 的表现介于显性纯合子 DD 和隐性纯合子 dd 的表现型之间，即在杂合子 Dd 中显性基因 D 和隐性基因 d 的作用均得到一定程度的表现。例如人类对苯硫脲(PTC)的尝味能力就是不完全显性遗传的典型性状。苯硫脲是一种白色结晶状物质，因含有 N—C ═S 结构而有苦涩味。有人能尝出其苦味，称 PTC 尝味者；有些人不能尝出其苦味，叫 PTC 味盲者。在我国汉族居民中，味盲者约占 1/10。

十五、不规则显性遗传

不规则显性(irregular dominance)遗传是杂合子的显性基因由于某种原因而不表现出相应的性状，因此在系谱中可以出现隔代遗传的现象。换言之，在具有某一显性基因的个体中，并不是每个个体都能表现出该显性基因所控制的性状。但是带有显性基因的某些个体，本身虽然不表现出显性性状，但他们却可以生出具有该性状的后代。显性基因不能表达的原因还不清楚，生物体的内外环境对基因表达所产生的影响和不同个体所具有的不同遗传背景可能是引起不规则显性的重要因素。多指(趾)症就是不规则显性的典型例子。

十六、共显性遗传

共显性(codominance)是一对等位基因之间，没有显性和隐性的区别，在杂合体时两种基因的作用都完全表现出来。例如人类的 ABO 血型、MN 血型和组织相容性抗原等的遗传属于这种遗传方式。

ABO 血型(OMIM110300)是一组复等位基因(A、B 和 O)所控制的。这一组复等位基因均位于第 9 号染色体上的 q34 位点。复等位基因(multiple alleles)来源于一个基因位点所发生的多次独立的突变，是基因突变多向性的表现。

十七、同一基因可产生显性或隐性突变

现已发现同一基因的不同突变可引起显性或隐性遗传病。如位于 11 号染色体短臂的 β 珠蛋白基因 127 位密码子的突变使 β 链的 127 位氨基酸从正常的谷氨酰胺变成了脯氨酸，从而形成 Hb Houston，导致 β^+-Houston-地中海贫血，其遗传方式为常染色体显性。而 β 珠蛋白基因 26 位密码子的突变，则使 β 链的 26 位氨基酸从正常的谷氨酸变成了赖氨酸，形成 Hb E，导致 β^+-E-地中海贫血，其遗传方式为常染色体隐性。类似的例子还有许多(表 5-6)。

了解上述经典遗传规律的例外情况，有助于我们辩证地认识问题、解决问题，最终揭示人类遗传病的奥秘。

表5-6 同一基因可产生显性或隐性突变的例子

遗传病名称	基因名称(符号及染色体定位)	常染色体显性病例	常染色体隐性病例
全身性甲状腺素抗性	甲状腺素受体-1(*THR1*;3p24.3)	全身性甲状腺素抗性(THRB,甘340精)	全身性甲状腺素抗性(*THRB*,外显子4-10缺失)
营养不良型大疱性表皮松懈	胶原蛋Ⅶ型(*COL7A1*;3p21.3)	显性营养不良型大疱性表皮松懈(*COL7A1*,甘2040丝)	隐性营养不良型大疱性表皮松懈(*COL7A1*,甲硫-赖)
联合垂体激素缺乏症	垂体特异性转录因子(*PIT1*;3p11)	联合垂体激素缺乏症(PIT1,精271色)	联合垂体激素缺乏症(*PIT1*,精172终止)
视网膜色素变性	视紫质(*RHO*;3q21-q24)	视网膜色素变性-4(视紫质相关视网膜色素变性,RHO,脯23组)	视网膜色素变性,常染色体隐性(*RHO*,内含子4给位,+1,G-T)
先天性肌强直	骨骼肌氯离子通道-1(*CLCN1*;7q35)	先天性肌强直,Thomsen型(CLCN1,甘230谷)	先天性肌强直,Becker型(CLCN1,苯丙413半胱)
β-地中海贫血	β珠蛋白(*HBB*;11p15.5)	β-Houston-地中海贫血(HBB,谷氨酰胺127脯)	β-E-地中海贫血(*HBB*,谷26赖)
Von Willebrand病	Von Willebrand因子(*VWF*;12p13.3)	Von Willebrand病Ⅰ型(VWF,精854谷氨酰胺)	Von Willebrand病因子Normardy-1(常染色体隐性血友病A;*VWF*,苏28甲硫)
Bernard-Soulier综合征	α血小板糖蛋白Ⅰb(*GP1BA*;17p)	常染色体显性Bernard-Soulier病(GP1BA,亮57苯丙)	Bernard-Soulier综合征(*GP1BA*,色343终止)
孤立型生长激素缺乏症	生长激素-1(*GH-1*;17q22-q24)	生长激素缺乏症(IGHDII;GH1,内含子4给位,+6,T-C)	生长激素缺乏症(*GH1*,2bp的缺失)
先天性慢通道肌无力综合征	尼古丁乙酰胆碱受体ε多肽(*CHRNE*;17号染色体)	先天性慢通道肌无力综合征;SCCMS(CHRNE,苏245脯)	先天性慢通道肌无力综合征;SCCMS(*CHRNE*,精64终止)
胰岛素抗性糖尿病伴黑棘皮病	胰岛素受体(*INSR*;19p13.2)	胰岛素抗性糖尿病伴黑棘皮病(INSR,甘996缬)	胰岛素抗性黑棘皮病(*INSR*,精735丝)

Summary

We define Mendelian diseases as diseases that are the result of a single mutant gene that has a large effect on phenotype and that are inherited in simple patterns similar to or identical with those described by Mendel for certain discrete characteristics in garden peas.Mendelian diseases are autosomal if they are encoded by genes on one of the 22 pairs of autosomes,and X-linked if encoded by a mutant gene on the X chromosome.Following Mendel,we define as dominant those conditions that are expressed in heterozygotes, *i.e.* ,individuals who have one copy of a mutant allele and one copy of a normal allele,and recessive those conditions that are clinically manifest only in individuals homozygous for the mutant allele,*i.e.* , carrying a double dose of an abnormal gene.It should be stressed that dominance and recessive refer to traits,or phenotypes,and not to genes.Although we sometimes speak of dominant and recessive genes,this is a shorthand and should be understood refering to traits.More than half are autosomal dominant traits,36% are autosomal recessive,and less than 5% are X-linked.

（左 伋）

参考文献

1. http://www.ncbi.nlm.nih.gov/omim
2. Gelehrter TD. 医学遗传学原理. 孙开来译. 北京:科学出版社,2001
3. Collins FS, McKusick VA. Implications of the human genome project for medical science. JAMA,2001,285(5):540-544
4. Williams SJ, Hayward NK. The impact of the human genome project on medical genetics. Trend Mol Med,2001,7(5):229-231
5. Nussbaum RL, McInnes RR, Willard HF. Genetics in Medicine. 7th ed. Elsevier(Singapore) Pte Ltd,2007

第六章　多基因遗传病

人类的一些遗传性状或某些遗传病的遗传基础不是一对基因，而是几对基因，每一对基因对遗传性状或遗传病形成的作用是微小的，故称为微效基因（minor gene）。但是，若干对基因作用积累之后，可以形成一个明显的表型效应，称为累加效应（additive effect）。这些基因也称累加基因（additive gene），因此这种性状或疾病的遗传方式称为多基因遗传（polygenic inheritance）或多因子遗传（multifactorial inheritance，MF）；同时由于上述的遗传性状或遗传病的发生，不仅取决于两个以上微效基因的累加作用，还受环境因子的影响，因此这类性状也称为复杂性状或复杂疾病（complex disease），例如，糖尿病、精神分裂症、哮喘、多发硬化症等。另一方面，目前的研究认为在多基因遗传中，除了微效基因外，可能存在一些起主要作用的所谓主基因（major gene），这使得多基因遗传更加复杂，但"主基因"对了解多基因疾病的发生、诊断、治疗和预防很有帮助。

第一节　数量性状的多基因遗传

传统上，多基因遗传的理论基础是多基因遗传的疾病或性状受许多的微效基因控制，这些微效基因彼此之间没有显性与隐性之分，是共显性的，有累加效应，这是它不同于单基因遗传之处。

一、数量性状与质量性状

多基因遗传性状或疾病与单基因遗传的性状或疾病有所不同，单基因遗传性状决定于一对基因。因此，其基因型和表型之间的相互关系比较直截了当，往往可以分出具有某性状（受累）和不具有某性状（未受累）两种不同类型，因此当该性状可以检测时，这一性状的变异在群体中的分布是不连续

的，可以明显地分为2～3群，所以单基因遗传的性状也称质量性状（qualitative character）。例如正常人血浆中苯丙氨酸羟化酶（PAH）的活性为100%，杂合携带者的PAH活性为正常人的45%～50%，苯丙酮尿症患者的PAH酶活性仅为正常人的0～5%。这分别决定于基因型PP、Pp、pp，若将此性状的变异作图，则可以看到三个峰（图6-1）。

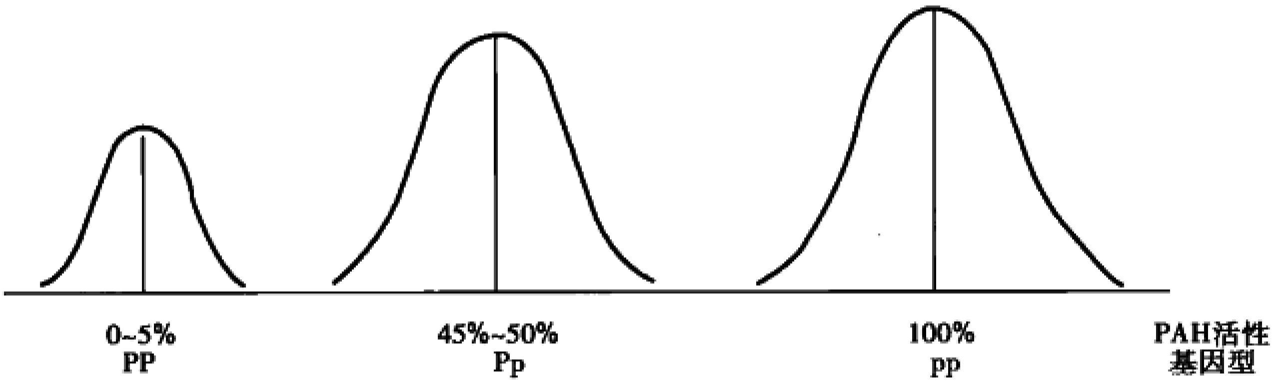


图6-1　质量性状变异分布图

多基因遗传性状的变异在群体中的分布是连续的，有一个峰，即平均值。不同个体间的差异只是量的变异，因此又称为数量性状（quantitative character）。例如，人的身高、智能、血压等。如果随机调查任何一个群体的身高，则极矮和极高的个体只占少数，大部分个体接近平均身高，而且呈现由矮向高逐渐过渡，将此身高变异分布绘成曲线，这种变异呈正态分布（图6-2）。

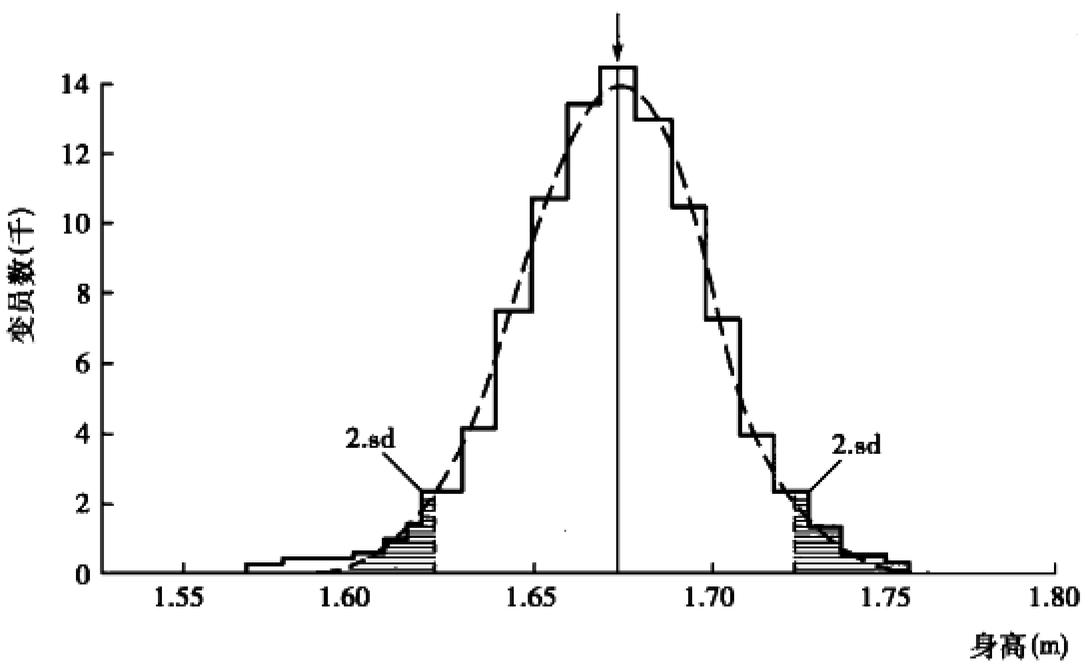


图6-2　数量性状（人身高）变异分布图

sd：标准差

二、数量性状的多基因遗传

数量性状是由许多数目不详、作用微小的等显性的微效基因控制的。那么，它是如何进行的呢？现以人的身高为例来解释数量性状的遗传机制。假设有三对非连锁的基因控制人类的身高，它们分别是AA′、BB′、CC′。这三对基因中A、B、C较A′、B′、C′对身高有增强作用，各可在平均身高（165cm）基础上增加5cm，因此基因型AABBCC个体为高身材个体（195cm）；而它们的等位基因A′、B′、C′则各在身高平均值的基础上减低5cm，故基因型A′A′B′B′C′C′个体为矮身材个体（135cm），介于这两者之间的基因取决于A、B、C和A′、B′、C′之间的组合，使人的身高从矮到高。假如亲代为一高身材（195cm）个体（AABBCC）与一矮身材（135cm）个体（A′A′B′B′C′C′）婚配，则子$_1$代将为杂合的基因型，即AA′BB′CC′，呈中等身材（165cm），但子$_1$代中也可能出现165cm左右的变异，这种变异完全是

环境因素作用的结果。假设相同基因型的子$_1$代个体间进行婚配，则这三对非连锁基因按分离律和自由组合律，可产生8种精子或卵子，精卵随机结合可产生64种基因型（表6-1），将各基因型按高矮数目分组，可以归并成7组：即6′0（表示有6个均带′的身高降低基因，0个不带′的身高增高基因）、5′1、4′2、3′3、2′4、1′5、0′6，它们的频数分布分别为1、6、15、20、15、6、1。再将这7组基因型组合频数分布做成柱形图，以横坐标为组合类型，纵坐标为频数，各柱形顶端连接成一线，即得到趋势近于正态分布的曲线（图6-3）。

表6-1 人身高三对基因遗传的基因组合

配子	ABC	A′BC	AB′C	ABC′	A′B′C	AB′C′	A′BC′	A′B′C′
ABC	AABBCC	AA′BBCC	AABB′CC	AABBCC′	AA′BB′CC	AABB′CC′	AA′BBCC′	AA′BB′CC′
A′BC	AA′BBCC	A′A′BBCC	AA′BB′CC	AA′BBCC′	A′A′BB′CC	AA′BB′CC′	A′A′BBCC′	A′A′BB′CC′
AB′C	AABB′CC	AA′BB′CC	AAB′B′CC	AABB′CC′	AA′B′B′CC	AAB′B′CC′	AA′BB′CC′	AA′B′B′CC′
ABC′	AABBCC′	AA′BBCC′	AABB′CC′	AABBC′C′	AA′BB′CC′	AABB′C′C′	AA′BBC′C′	AA′BB′C′C′
A′B′C	AA′BB′CC	A′A′BB′CC	AA′B′B′CC	AA′BB′CC′	A′A′B′B′CC	AA′B′B′CC′	A′A′BB′CC′	A′A′B′B′CC′
AB′C′	AABB′CC′	AA′BB′CC′	AAB′B′CC′	AABB′C′C′	AA′B′B′CC′	AAB′B′C′C′	AA′BB′C′C′	AA′B′B′C′C′
A′BC′	AA′BBCC′	A′A′BBCC′	AA′BB′CC′	AA′BBC′C′	A′A′BB′CC′	AA′BB′C′C′	A′A′BBC′C′	A′A′BB′C′C′
A′B′C′	AA′BB′CC′	A′A′BB′CC′	AA′B′B′CC′	AA′BB′C′C′	A′A′B′B′CC′	AA′B′B′C′C′	A′A′BB′C′C′	A′A′B′B′C′C′

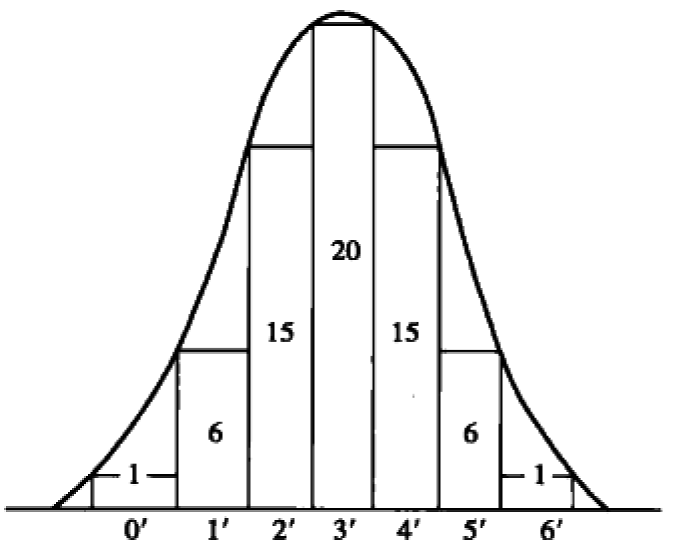


图6-3 子$_1$代身高变异分布图

当然，决定身高及其他数量性状的基因远不止3对，而且许多研究也显示每一个基因的作用也并不是等同的，还有所谓的主基因（在数量性状的形成过程起主要作用），加上环境因素的影响，因此数量性状在群体中的分布就更为复杂，通常形成一种连续的正态分布曲线。

人的身高除受遗传因素影响外，还受到各种环境因素的影响，如营养好坏、阳光充足与否、是否进行体育锻炼等。因此，环境因素对表现型有重要作用，它们对某种性状的产生起着增强或抑制作用。

人的身高取决于多对微效基因的组合。因此，双亲的身高决定着子女身高。但会发现身材高大的双亲，其子女虽然身高仍偏高，但多数将比他们的父母要矮，即比父母更接近于人群的平均身高；两个极矮的双亲，其子女身高高于父母身高平均值。这种现象早在1926年由英国著名的科学家Galton发现，并称之为“平均值的回归”。他通过测量204对双亲和他们的928名成年女子身高获此结论：如果双亲身高平均值高于群体平均值，子女平均值就低于其双亲平均值，但接近群体身高平均值；如果双亲身高平均值低于群体平均值，则子女身高高于其双亲平均值，但接近群体身高平均值。这就是说，数量性状在遗传过程中子代将向人群的平均值靠拢，这就是回归现象。这种现象也表现于其他相似的数量性状。

多基因遗传性状或数量性状表达中所反映的回归现象，对理解多基因遗传病遗传特点有着重要指导意义。

第二节 多基因病的遗传

多基因病是一类患病率较高、发病较为复杂的疾病。在分析和研究其病因、发病机制、再发风险估计等,不仅要分析遗传因素,同时又不能忽视环境因素影响。

一、易患性与发病阈值

在多基因遗传病发生中,遗传因素和环境因素共同作用决定一个个体患某种遗传病的可能性称为易患性(liability)。一般群体中,易患性很高或很低的个体都很少,大部分个体都接近平均值。因此,群体中的易患性变异也呈正态分布。易感性(susceptibility)特指由遗传因素决定的患病风险,仅代表个体所含有的遗传因素;但在一定的环境条件下,易感性高低可代表易患性高低。当一个个体易患性高到一定限度就可能发病。这种由易患性所导致的多基因遗传病发病最低限度称为发病阈值(threshold)。这样,阈值将群体中连续分布的易患性变异分为两部分,即一部分是正常群体,另一部分是患病群体(图6-4)。阈值标志着在一定的环境条件下,发病所必需的最低的致病基因数量,所以多基因遗传性状亦属于阈值性状。

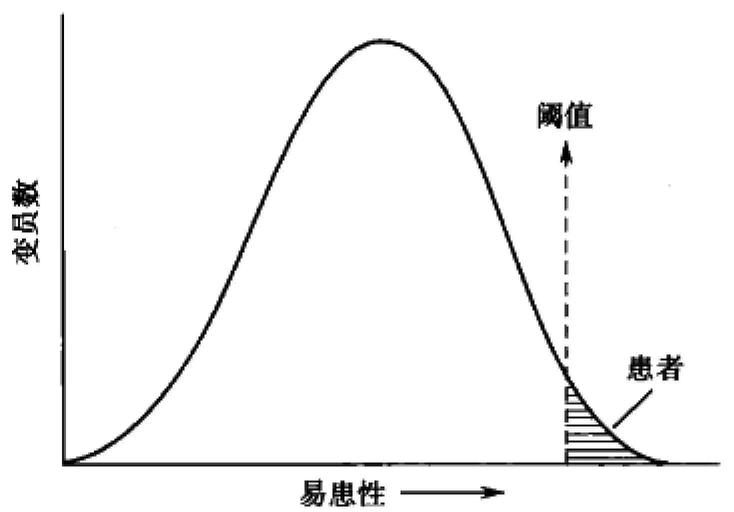


图6-4 群体易患性变异分布图

一个个体的易患性高低无法测量,但是,一个群体的易患性平均值可以从该群体的患病率做出估计。利用正态分布平均值(或均值μ)与标准差(σ)之间已知关系,可由患病率估计群体的发病阈值与易患性平均值之间的距离,这距离是以正态分布的标准差作为衡量单位。根据正态分布曲线下的总面积为100%,可推算得到均数加减任何数量标准差的范围内,曲线与横轴之间所包括面积占曲线下的总面积的比例。多基因遗传病的群体易患性呈正态分布,因此,它必然具有正态分布的特征,从图6-5中可以得到以下关系:①$\mu \pm 1\sigma$(以平均值μ为0,左右1个标准差)范围内的面积占正态分布曲线下的总面积的68.28%,此范围以外的面积占31.72%,左右侧各占约16%;②$\mu \pm 2\sigma$范围内的面积占正态分布曲线下的总面积的95.46%,此范围以外的面积占4.54%,左右侧各占约2.3%;③$\mu \pm 3\sigma$范围内的面积占正态分布曲线下的总面积的99.74%,此范围以外的面积占0.26%,左右侧各占约0.13%。

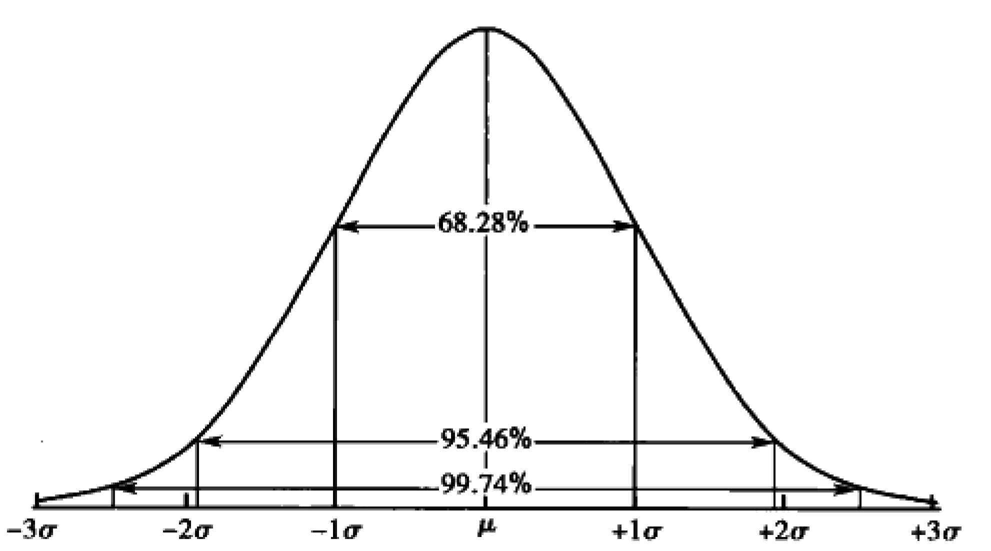


图6-5 正态分布曲线中μ与σ的关系

多基因遗传病易患性正态分布曲线下的面积代表总人群,其易患性超过阈值的那部分面积为患者所占的百分数,即患病率。所以人群中某一种多基因遗传病的患病率即为超过阈值的那部分面积。从其患病率就可以得出阈值距离均数有几个标准差,这只要查阅正态分布表即可(表6-2)。易患性正态分布曲线右侧尾部的面积代表患病率。例如,冠心病的群体患病率为2.3%～2.5%,其阈值与易患性平均值距离约2σ;而先天性畸形足的群体患病率仅为0.13%,其阈值与易患性平均值距离约3σ。

表6-2 正态分布的X和a值表

q%	X	a	q%	X	a	q%	X	a	q%	X	a	q%	X	a
0.01	3.719	3.960	0.31	2.737	3.040	0.61	2.506	2.829	0.91	2.361	2.697	1.21	2.254	2.600
0.02	3.540	3.790	0.32	2.727	3.030	0.62	2.501	2.823	0.92	2.357	2.693	1.22	2.251	2.597
0.03	3.432	3.687	0.33	2.716	3.021	0.63	2.495	2.818	0.93	2.353	2.690	1.23	2.248	2.594
0.04	3.353	3.613	0.34	2.706	3.012	0.64	2.489	2.813	0.94	2.349	2.686	1.24	2.244	2.591
0.05	3.291	3.554	0.35	2.697	3.003	0.65	2.484	2.808	0.95	2.346	2.683	1.25	2.241	2.589
0.06	3.239	3.507	0.36	2.687	2.994	0.66	2.478	2.803	0.96	2.342	2.679	1.26	2.238	2.586
0.07	3.195	3.464	0.37	2.678	2.986	0.67	2.473	2.798	0.97	2.338	2.676	1.27	2.235	2.583
0.08	3.156	3.429	0.38	2.669	2.978	0.68	2.468	2.793	0.98	2.334	2.672	1.28	2.232	2.580
0.09	3.121	3.397	0.39	2.661	2.969	0.69	2.462	2.798	0.99	2.330	2.669	1.29	2.229	2.578
0.10	3.090	3.367	0.40	2.652	2.962	0.70	2.457	2.784	1.00	2.326	2.665	1.30	2.226	2.575
0.11	3.062	3.341	0.41	2.644	2.954	0.71	2.452	2.779	1.01	2.323	2.662	1.31	2.223	2.572
0.12	3.036	3.317	0.42	2.636	2.947	0.72	2.447	2.775	1.02	2.319	2.658	1.32	2.220	2.570
0.13	3.012	3.294	0.43	2.628	2.939	0.73	2.442	2.770	1.03	2.315	2.655	1.33	2.217	2.567
0.14	2.989	3.273	0.44	2.620	2.932	0.74	2.437	2.766	1.04	2.312	2.652	1.34	2.214	2.564
0.15	2.968	3.258	0.45	2.612	2.925	0.75	2.432	2.761	1.05	2.308	2.649	1.35	2.211	2.562
0.16	2.948	3.234	0.46	2.605	2.918	0.76	2.428	2.757	1.06	2.304	2.645	1.36	2.209	2.559
0.17	2.929	3.217	0.47	2.597	2.911	0.77	2.423	2.753	1.07	2.301	2.642	1.37	2.206	2.557
0.18	2.911	3.201	0.48	2.590	2.905	0.78	2.418	2.748	1.08	2.297	2.639	1.38	2.203	2.554
0.19	2.894	3.185	0.49	2.583	2.898	0.79	2.414	2.744	1.09	2.294	2.636	1.39	2.200	2.552
0.20	2.878	3.170	0.50	2.576	2.892	0.80	2.409	2.740	1.10	2.290	2.633	1.40	2.197	2.549
0.21	2.863	3.156	0.51	2.569	2.886	0.81	2.404	2.736	1.11	2.287	2.630	1.41	2.194	2.547
0.22	2.848	3.142	0.52	2.562	2.880	0.82	2.400	2.732	1.12	2.283	2.627	1.42	2.192	2.544
0.23	2.843	3.129	0.53	2.556	2.873	0.83	2.395	2.728	1.13	2.280	2.624	1.43	2.189	2.542
0.24	2.820	3.117	0.54	2.549	2.868	0.84	2.391	2.724	1.14	2.277	2.621	1.44	2.186	2.539
0.25	2.807	3.104	0.55	2.543	2.862	0.85	2.387	2.720	1.15	2.273	2.618	1.45	2.183	2.537
0.26	2.794	3.093	0.56	2.536	2.856	0.86	2.382	2.716	1.16	2.270	2.616	1.46	2.181	2.534
0.27	2.782	3.081	0.57	2.530	2.850	0.87	2.378	2.712	1.17	2.267	2.612	1.47	2.178	2.532
0.28	2.770	3.070	0.58	2.524	2.845	0.88	2.374	2.708	1.18	2.264	2.609	1.48	2.175	2.529
0.29	2.759	3.060	0.59	2.518	2.839	0.89	2.370	2.704	1.19	2.260	2.606	1.49	2.173	2.527
0.30	2.748	3.050	0.60	2.512	2.834	0.90	2.366	2.701	1.20	2.257	2.603	1.50	2.170	2.525

续表

q%	X	a	q%	X	a	q%	X	a	q%	X	a	q%	X	a
1.51	2.167	2.522	1.86	2.084	2.447	4.1	1.739	2.144	7.6	1.433	1.881	11.1	1.221	1.705
1.52	2.165	2.520	1.87	2.081	2.445	4.2	1.728	2.135	7.7	1.426	1.876	11.2	1.216	1.701
1.53	2.162	2.518	1.88	2.079	2.444	4.3	1.717	2.125	7.8	1.419	1.870	11.3	1.211	1.696
1.54	2.160	2.515	1.89	2.077	2.442	4.4	1.706	2.116	7.9	1.412	1.864	11.4	1.206	1.692
1.55	2.157	2.513	1.90	2.075	2.440	4.5	1.695	2.106	8.0	1.405	1.858	11.5	1.200	1.688
1.56	2.155	2.511	1.91	2.073	2.438	4.6	1.685	2.097	8.1	1.398	1.853	11.6	1.195	1.684
1.57	2.152	2.508	1.92	2.071	2.436	4.7	1.675	2.088	8.2	1.392	1.847	11.7	1.190	1.679
1.58	2.149	2.506	1.93	2.068	2.434	4.8	1.665	2.080	8.3	1.385	1.842	11.8	1.185	1.675
1.59	2.147	2.504	1.94	2.066	2.432	4.9	1.655	2.071	8.4	1.379	1.836	11.9	1.180	1.671
1.60	2.144	2.502	1.95	2.064	2.430	5.0	1.645	2.063	8.5	1.372	1.831	12.0	1.175	1.667
1.61	2.142	2.499	1.96	2.062	2.428	5.1	1.635	2.054	8.6	1.366	1.825	12.1	1.170	1.663
1.62	2.139	2.497	1.97	2.060	2.426	5.2	1.626	2.046	8.7	1.359	1.820	12.2	1.165	1.659
1.63	2.137	2.495	1.98	2.058	2.425	5.3	1.616	2.038	8.8	1.353	1.815	12.3	1.160	1.655
1.64	2.135	2.493	1.99	2.056	2.423	5.4	1.607	2.030	8.9	1.347	1.810	12.4	1.155	1.651
1.65	2.132	2.491	2.0	2.054	2.421	5.5	1.598	2.023	9.0	1.341	1.804	12.5	1.150	1.647
1.66	2.130	2.489	2.1	2.034	2.403	5.6	1.589	2.015	9.1	1.335	1.799	12.6	1.146	1.643
1.67	2.127	2.486	2.2	2.014	2.386	5.7	1.580	2.007	9.2	1.329	1.794	12.7	1.141	1.639
1.68	2.125	2.484	2.3	1.995	2.369	5.8	1.572	2.000	9.3	1.323	1.789	12.8	1.136	1.635
1.69	2.122	2.482	2.4	1.977	2.353	5.9	1.565	1.993	9.4	1.317	1.784	12.9	1.131	1.631
1.70	2.120	2.480	2.5	1.960	2.338	6.0	1.555	1.985	9.5	1.311	1779	13.0	1.126	1.627
1.71	2.118	2.478	2.6	1.943	2.323	6.1	1.546	1.978	9.6	1.305	1.774	13.1	1.122	1.623
1.72	2.115	2.476	2.7	1.927	2.309	6.2	1.538	1.971	9.7	1.299	1.769	13.2	1.117	1.620
1.73	2.113	2.474	2.8	1.911	2.295	6.3	1.530	1.964	9.8	1.239	1.765	13.3	1.112	1.616
1.74	2.111	2.472	2.9	1.896	2.281	6.4	1.522	1.957	9.9	1.287	1.760	13.4	1.108	1.612
1.75	2.108	2.470	3.0	1.881	2.268	6.5	1.514	1.951	10.0	1.282	1.755	13.5	1.103	1.608
1.76	2.106	2.467	3.1	1.866	2.255	6.6	1.506	1.944	10.1	1.276	1.750	13.6	1.098	1.606
1.77	2.104	2.465	3.2	1.852	2.243	6.7	1.499	1.937	10.2	1.270	1.746	13.7	1.094	1.601
1.78	2.101	2.463	3.3	1.838	2.231	6.8	1.491	1.931	10.3	1.265	1.741	13.8	1.089	1.597
1.79	2.099	2.461	3.4	1.825	2.219	6.9	1.483	1.924	10.4	1.259	1.736	13.9	1.085	1.593
1.80	2.097	2.459	3.5	1.812	2.208	7.0	1.476	1.918	10.5	1.254	1.732	14.0	1.080	1.590
1.81	2.095	2.457	3.6	1.799	2.197	7.1	1.468	1.912	10.6	1.248	1.727	14.1	1.076	1.586
1.82	2.092	2.455	3.7	1.787	2.186	7.2	1.461	1.906	10.7	1.243	1.723	14.2	1.071	1.583
1.83	2.090	2.453	3.8	1.774	2.175	7.3	1.454	1.899	10.8	1.237	1.718	14.3	1.067	1.579
1.84	2.088	2.451	3.9	1.762	2.165	7.4	1.447	1.893	10.9	1.232	1.714	14.4	1.063	1.575
1.85	2.086	2.449	4.0	1.751	2.154	7.5	1.440	1.887	11.0	1.227	1.709	14.5	1.058	1.572

续表

q%	X	a	q%	X	a	q%	X	a	q%	X	a	q%	X	a
14.6	1.054	1.568	16.5	0.974	1.504	18.4	0.900	1.446	20.3	0.831	1.391	33.0	0.440	1.097
14.7	1.049	1.565	16.6	0.970	1.501	18.5	0.896	1.443	20.4	0.827	1.389	34.0	0.412	1.075
14.8	1.045	1.561	16.7	0.966	1.498	18.6	0.893	1.440	20.5	0.824	1.386	35.0	0.385	1.058
14.9	1.041	1.558	16.8	0.962	1.495	18.7	0.889	1.437	20.6	0.820	1.383	36.0	0.358	1.039
15.0	1.036	1.554	16.9	0.958	1.492	18.8	0.885	1.434	20.7	0.817	1.381	37.0	0.332	1.020
15.1	1.032	1.551	17.0	0.954	1.489	18.9	0.882	1.431	20.8	0.813	1.378	38.0	0.305	1.002
15.2	1.028	1.548	17.1	0.950	1.485	19.0	0.878	1.428	20.9	0.810	1.375	39.0	0.279	0.984
15.3	1.024	1.544	17.2	0.946	1.482	19.1	0.874	1.425	21.0	0.806	1.372	40.0	0.253	0.966
15.4	1.019	1.541	17.3	0.942	1.479	19.2	0.871	1.422	22.0	0.772	1.346	41.0	0.228	0.948
15.5	1.015	1.537	17.4	0.938	1.476	19.3	0.867	1.420	23.0	0.739	1.320	42.0	0.202	0.931
15.6	1.011	1.534	17.5	0.935	1.473	19.4	0.863	1.417	24.0	0.706	1.295	43.0	0.176	0.913
15.7	1.007	1.531	17.6	0.931	1.470	19.5	0.860	1.414	25.0	0.674	1.271	44.0	0.151	0.896
15.8	1.003	1.527	17.7	0.927	1.467	19.6	0.856	1.411	26.0	0.643	1.248	45.0	0.126	0.880
15.9	0.999	1.524	17.8	0.923	1.464	19.7	0.852	1.408	27.0	0.613	1.225	46.0	0.100	0.863
16.0	0.994	1.521	17.9	0.919	1.461	19.8	0.849	1.405	28.0	0.583	1.202	47.0	0.075	0.846
16.1	0.990	1.517	18.0	0.915	1.458	19.9	0.845	1.403	29.0	0.553	1.180	48.0	0.050	0.830
16.2	0.986	1.514	18.1	0.912	1.455	20.0	0.842	1.400	30.0	0.524	1.159	49.0	0.025	0.814
16.3	0.982	1.511	18.2	0.908	1.452	20.1	0.838	1.397	31.0	0.496	1.138	50.0	0.000	0.798
16.4	0.978	1.508	18.3	0.904	1.449	20.2	0.834	1.394	32.0	0.468	1.118			

可见，一种多基因病的易患性的平均值与阈值越近，表明易患性高阈值低，群体患病率高；相反，易患性的平均值与阈值越远，表明易患性低阈值高，群体患病率低（图6-6）。

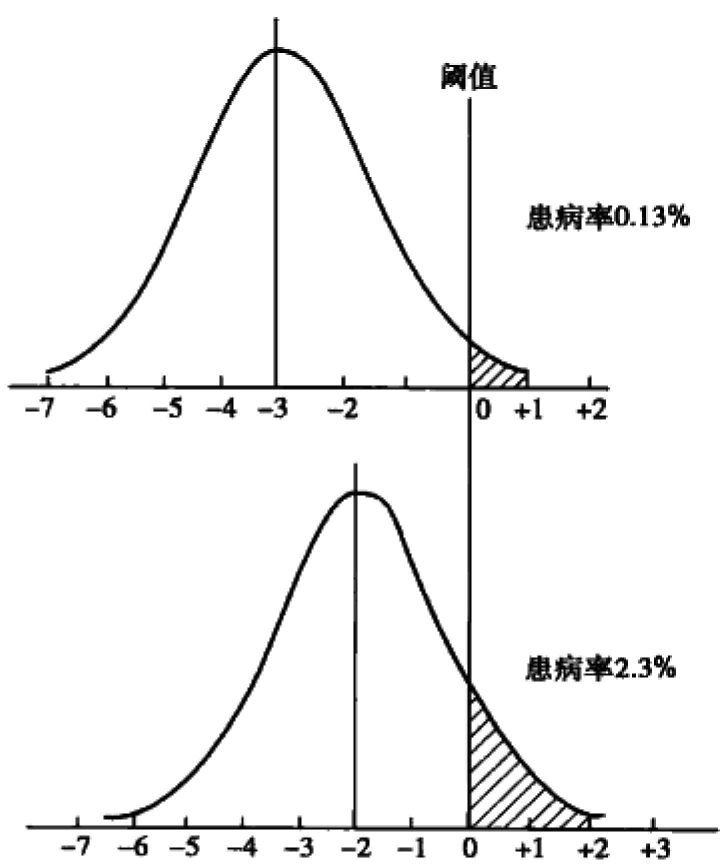


图6-6 易患性的平均值和阈值距离与患病率的关系

二、遗传度

已知表型是基因型与环境共同作用的结果，因此，一个数量性状的变异总是包括遗传变异和环境变异两个组成部分。根据这一概念，可用方差来表示各项变异的程度，即：

$$表型方差(V_P)=遗传方差(V_G)+环境方差(V_E)$$

遗传方差就是基因型方差，属于可遗传的变异；而环境方差属于不可遗传的变异。为了衡量遗传方差和环境方差在表型方差中的相对作用，常采用遗传度这一指标。遗传度是指遗传方差占表型方差的比值，它表示数量性状从亲代传递给子代的相对能力。一般用 h^2 表示，其公式为：

$$h^2 = \frac{V_G}{V_P} \times 100\% \tag{6-1}$$

该公式求得的遗传度为广义遗传度。其数值愈大，表明这一性状变异受遗传的影响愈大，受环境的影响愈小；其数值愈小，表明这一性状变异受遗传的影响愈小，受环境的影响愈大。

如进一步分析，遗传方差还可分为三个组成部分，即：

遗传方差(V_G) = 加性方差(V_A) + 显性方差(V_D) + 上位性方差(V_I)

在该公式中，V_A 是多基因累加效应的方差，是可固定遗传的部分；V_D 是基因显性效应的方差，将随世代的增进而消失，是不能固定遗传的部分；而 V_I 和 V_E 同样都是不能遗传的。因此，狭义遗传度是指加性方差占表型方差的比值，即：

$$h^2 = \frac{V_A}{V_P} \times 100\% = \frac{V_A}{(V_A + V_D + V_I) + V_E} \times 100\% \tag{6-2}$$

由于需要分别估算 V_A、V_D 和 V_I，故计算较繁，但狭义遗传度可以更确切地估算可遗传的变异对表型的影响。因为 V_A 一定小于 V_G，所以狭义遗传度总是小于广义遗传度。通常所谓的遗传度是指狭义遗传度。

在多基因遗传病中，遗传度(heritability)的含义是多基因累加效应对疾病易患性变异的贡献大小。遗传度愈大，表明遗传因素对病因的贡献愈大。如果一种疾病其易患性变异全由遗传因素所决定，遗传度就是100%，这种情况是极少见的。在遗传度高的疾病中，遗传度可高达70% ~80%，这表明遗传因素在决定疾病易患性变异上有重要作用，环境因素的作用较小；在遗传度低的疾病中，遗传度仅为30% ~40%，这表明在决定疾病易患性变异上，环境因素有重要作用，而遗传因素的作用不显著，就不会出现明显的家族聚集现象。

计算多基因遗传病遗传度的高低在临床实践上有重要意义，其计算方法有两种：

(一) Falconer 公式

Falconer 公式是根据先证者亲属的发病率与遗传度有关而建立的。亲属发病率越高，遗传度越大，所以可通过调查先证者亲属发病率和一般人群的发病率，算出遗传度(h^2 或 H)。

$$h^2 = b/k \tag{6-3}$$

已知一般人群的发病率时，用下式计算回归系数 b 及其方差 V_b：

$$b = \frac{X_g - X_r}{a_g} \tag{6-4}$$

$$V_b = \left(\frac{1}{a_g}\right)^2 W_r = \left(\frac{1}{a_g}\right)^2 \left(\frac{1-q_r}{a_r^2 A_r}\right) \tag{6-5}$$

缺乏一般人群的发病率时，可设立对照组，调查对照组亲属的发病率，用下式计算回归系数 b 及其方差 V_b：

$$b = \frac{p_c(X_c - X_r)}{a_c} \tag{6-6}$$

$$V_b = \left(\frac{p_c}{a_c}\right)^2 W_r = \left(\frac{p_c}{a_c}\right)^2 \left(\frac{1-q_r}{a_r^2 A_r}\right) \tag{6-7}$$

$$S_{h^2} = \frac{\sqrt{V_b}}{k} \tag{6-8}$$

$$t = \frac{h^2}{S_{h^2}} \tag{6-9}$$

在估算遗传度时，如果只估算一级亲属的遗传度，可能会因为一级亲属常共同生活，暴露于相同的环境，而遗传度正是根据亲属间的相似程度来估算的，共同的环境暴露也可导致亲属间的相似程度增高，从而使估算的遗传度产生偏倚，因此，有必要同时估算二级亲属和三级亲属的遗传度，然后计算遗传度的加权平均值和遗传度标准误的加权平均值。但是，对二级亲属和三级亲属的遗传度也要进行显著性检验，仅在其遗传度存在统计学意义时，才可求其加权平均值。

$$h^2\text{ 的加权平均值} = \frac{\frac{h_1^2}{S_1^2} + \frac{h_2^2}{S_2^2} + \frac{h_3^2}{S_3^2} + \cdots + \frac{h_n^2}{S_n^2}}{\frac{1}{S_1^2} + \frac{1}{S_2^2} + \frac{1}{S_3^2} + \cdots + \frac{1}{S_n^2}} \tag{6-10}$$

$$S_{h^2}\text{的加权平均值} = \frac{1}{\sqrt{\frac{1}{S_1^2} + \frac{1}{S_2^2} + \frac{1}{S_3^2} + \cdots + \frac{1}{S_n^2}}} \tag{6-11}$$

b—亲属易患性对先证者易患性的回归系数

k—亲属系数（一级亲属—指一个人与其双亲、子女和同胞之间，其基因有 1/2 的可能性是相同的；二级亲属—指一个人与其叔、伯、姑、舅、姨、祖父母和外祖父母之间，其基因有 1/4 的可能性是相同的；三级亲属—指一个人与其表兄妹、堂兄妹、曾祖父母之间，其基因有 1/8 的可能性是相同的）

X_g——一般群体易患性平均值与阈值之间的标准差数

X_c—对照组亲属中的易患性平均值与阈值之间的标准差数

X_r—先证者亲属易患性平均值与阈值之间的标准差数

a_g——一般群体易患性平均值与一般群体中患者易患性平均值之间的标准差数

a_c—对照组亲属易患性平均值与对照组亲属中患者易患性平均值之间的标准差数

a_r—先证者亲属易患性平均值与先证者亲属中患者易患性平均值之间的标准差数

q_g——一般群体发病率

q_c—对照亲属发病率，$p_c = 1 - q_c$

q_r—先证者亲属发病率

A_r—先证者亲属中的病例数

V_b—b 的方差

S_{h^2}—h^2 的标准误

h_1^2、h_2^2、h_3^2、……、h_n^2—分别为各级亲属的遗传度

S_1、S_2、S_3、……、S_n—分别为各级亲属遗传度的标准误

X_g、X_c、X_r 和 a_g、a_c、a_r 均可由一般群体发病率、对照亲属发病率和先证者亲属发病率查 Falconer 表（表 6-2）得到。

在研究实践中，作者发现有人对亲属的级别难以作出正确判断，这里根据作者的体会介绍一种判断亲属级别的方法。首先明确以先证者为核心，先证者的父母、同胞和子女是先证者的一级亲属，这些一级亲属的一级亲属是先证者的二级亲属，这些二级亲属的一级亲属是先证者的三级亲属，依此类推。但要注意每个家系成员都只能考虑一次，不能重复考虑。如先证者本人也是其一级亲属的一级亲属，如重复考虑，先证者将变成其本人的二级亲属。

例如，有人调查先天性房间隔缺损在一般群体中的发病率为 1/1000（0.1%），在 100 个先证者的家系中调查，先证者的一级亲属共有 669 人（双亲 200 人，同胞 279 人，子女 190 人），其中有 22 人（A_r）发病，依次求得先证者一级亲属的发病率为 22/669 × 100% = 3.3%（q_r），然后查 Falconer 表。按

群体发病率查得 X_g 和 a_g，再根据亲属发病率查得 X_r 和 a_r，然后代入公式求出 b 值。

$$b=\frac{X_g-X_r}{a_g}=\frac{3.090-1.838}{3.367}=0.37$$

将 b 值代入公式：

$$h^2=b/k=0.37/0.5=0.74=74\%$$

$$V_b=\left(\frac{1}{a_g}\right)^2\left(\frac{1-q_r}{a_r^2A_r}\right)=\left(\frac{1}{3.367}\right)^2\left(\frac{1-0.033}{2.231^2\times22}\right)=0.000778965$$

$$S_{h^2}=\frac{\sqrt{V_b}}{k}=\frac{\sqrt{0.000778965}}{0.5}=0.0558$$

$$t=\frac{h^2}{S_{h^2}}=\frac{0.74}{0.0558}=13.26$$

$$P<0.01$$

以上计算结果表明，遗传因素对先天性房间隔缺损发生的贡献为74%，经显著性检验该遗传度有统计学意义。

在缺乏一般人群发病率数据时，可选择与病例组匹配的对照组，调查对照组亲属的发病率，用先证者亲属和对照亲属的发病率计算遗传度。

再如，对江苏启东肝癌的调查发现，肝癌患者一级亲属6591人中，有359人(A_r)发病，其发病率为5.45%(q_r)；在年龄和性别均与患者相应的无病对照者的5227名一级亲属中，有54人患肝癌，发病率 $q_c=0.0103=1.03\%$。$p_c=1-q_c=0.9897$，分别查得 X_r、X_c 和 a_r、a_c，然后代入公式求出 b 值。

$$b=\frac{p_c(X_c-X_r)}{a_c}=\frac{0.9897(2.315-1.603)}{2.655}=0.2654$$

将 b 值代入公式：

$$h^2=b/k=0.2654/0.5=0.5308=53.08\%$$

$$V_b=\left(\frac{p_c}{a_c}\right)^2\left(\frac{1-q_r}{a_r^2A_r}\right)=\left(\frac{0.9897}{2.655}\right)^2\left(\frac{1-0.0545}{2.0265^2\times359}\right)=0.000087559$$

$$S_{h^2}=\frac{\sqrt{V_b}}{k}=\frac{\sqrt{0.000087559}}{0.5}=0.0187$$

$$t=\frac{h^2}{S_{h^2}}=\frac{0.5308}{0.0187}=28.39$$

$$P<0.01$$

以上计算结果表明，遗传因素对肝癌发生的贡献超过50%，经显著性检验该遗传度有统计学意义。

在这里，我们的两个实例仅计算了一级亲属的遗传度，二级和三级亲属遗传度的计算方法与一级亲属相同，但应注意的是，不同级别亲属遗传度的计算均是分别进行的，不同级别的亲属，其 b、V_b 和 k 的数值是不同的。我们计算的遗传度是一个点估计值，对其也可计算可信区间，通常计算其95%可信区间，即 $h^2\pm1.96S_{h^2}$。

（二）Holzinger 公式

Holzinger(1929)公式是根据遗传度越高的疾病，一卵双生的患病一致率与二卵双生患病一致率相差越大而建立的。

所谓一卵双生(monozygotic twin)是指由一个受精卵形成的两个双生子，他们的遗传基础相同，发育环境则可能存在差异；二卵双生(dizygotic twin)是指由两个受精卵形成的两个双生子，他们的遗传基础不同(其差异程度与一般同胞间相同)，发育环境也可能存在差异。所谓患病一致率是指双生子

中一个患某种疾病，另一个也患同样疾病的频率。

$$h^2 = \frac{C_{MZ} - C_{DZ}}{100 - C_{DZ}} \tag{6-12}$$

$$S_{h^2} = \sqrt{\left[\frac{1 - C_{MZ}}{(1 - C_{DZ})^2}\right]^2 \cdot \frac{C_{DZ}(1 - C_{DZ})}{n_2} + \left(\frac{1}{1 - C_{DZ}}\right)^2 \cdot \frac{C_{MZ}(1 - C_{MZ})}{n_1}} \tag{6-13}$$

C_{MZ}——一卵双生子的同病率

C_{DZ}—二卵双生子的同病率

n_1——一卵双生子的对子数

n_2—二卵双生子的对子数

例如，对躁狂抑郁性精神病的调查表明，在 15 对一卵双生子中，共同患病的有 10 对；在 40 对二卵双生子中，共同患病的有 2 对。依此来计算一卵双生子的同病率为 67%，二卵双生子的同病率为 5%。代入上式：

$$h^2 = \frac{C_{MZ} - C_{DZ}}{100 - C_{DZ}} = \frac{67 - 5}{100 - 5} = 0.65 = 65\%$$

$$S_{h^2} = \sqrt{\left[\frac{1 - 0.67}{(1 - 0.05)^2}\right]^2 \cdot \frac{0.05(1 - 0.05)}{40} + \left(\frac{1}{1 - 0.05}\right)^2 \cdot \frac{0.67(1 - 0.67)}{15}} = 0.1284$$

以上结果表明，在躁狂抑郁性精神病中，遗传因素的贡献为 65%。

一些常见的多基因遗传病的患病率和遗传度见表 6-3。

表 6-3 常见多基因遗传病的群体患病率、先证者一级亲属患病率、性别比和遗传度

疾病	一般群体患病率(%)	患者一级亲属患病率(%)	男/女	遗传度(%)
原发性高血压	4～8	20～30	1	62
哮喘	4	20	0.8	80
消化性溃疡	4	8	1	37
冠心病	2.5	7	1.5	65
精神分裂症	1.0	10	1	80
糖尿病(早发型)	0.2	2～5	1	75
脊柱裂	0.3	4	0.8	60
无脑儿	0.2	2	0.4	60
唇裂±腭裂	0.17	4	1.6	76
腭裂	0.04	2	0.7	76
先天性畸形足	0.1	3	2.0	68
先天性髋关节脱位	0.07	4	0.2	70
先天性幽门狭窄	0.3	男先证者 2	5.0	75
		女先证者 10		
先天性巨结肠	0.02	男先证者 2	4.0	80
		女先证者 8		
强直性脊椎炎	0.2	男先证者 7	0.2	70
		女先证者 2		

应当指出，遗传度估计值是由特定环境中特定人群的患病率估算得到的，不宜外推到其他人群和其他环境；同时，遗传度是群体统计量，用到个体毫无意义。如果某种疾病的遗传度为50%，不能说某个患者的发病一半由遗传因素决定，一半由环境因素决定，而应该说在这种疾病的总变异中，一半与遗传变异有关，一半与环境变异有关。遗传度的估算仅适合于没有遗传异质性，而且也没有主基因效应的疾病。若导致疾病的多基因中有一个显性主基因，那么估算的遗传度可以超过100%；若主基因为隐性基因，则由先证者的同胞估算的遗传度可以高于由父母或子女估算的遗传度。因此，只有当由同胞、父母和子女分别估算的遗传度相近似时，这个遗传度才是合适的。同时也才能认为该疾病的发生可能是多基因遗传的结果。

三、影响多基因遗传病再发风险估计的因素

（一）患病率与亲属级别有关

多基因遗传病发病有明显的家族聚集倾向，患者亲属患病率高于群体患病率，而且随着与患者亲缘关系级别的变远（或亲缘系数增大）患病率递减，向群体患病率靠拢。这一点与 Galton 提出的数量性状在亲属中存在回归现象相一致（表6-4）。

表6-4　多基因遗传病中亲属级别和患病率之间的关系

人　群	马蹄内翻足	唇裂±腭裂	先天性髋关节脱位（女）	先天性幽门狭窄（男）
一般群体	0.001	0.001	0.002	0.005
单卵双生	0.3（×300）	0.4（×400）	0.4（×200）	0.4（×80）
一级亲属	0.025（×25）	0.04（×40）	0.05（×25）	0.05（×10）
二级亲属	0.005（×5）	0.007（×7）	0.006（×3）	0.025（×5）
三级亲属	0.002（×2）	0.003（×3）	0.004（×2）	0.0075（×1.5）

在相当多的多基因遗传病中，群体患病率 q_g 常在0.1%～1%，遗传度为70%～80%之间，那么患者一级亲属的再患风险可利用 Edwards（1960）公式，其下标内容为患者一级亲属再发风险，q_r 是群体患病率 q_g 的平方根，即 $q_r = \sqrt{q_g}$；当遗传度低于70%～80%时，患者一级亲属再发风险低于群体患病率的平方根；当遗传度高于70%～80%时，一级亲属再患风险高于群体患病率的平方根。例如：唇裂的群体患病率为0.17%，其遗传度为76%，患者一级亲属再发风险 $q_r = \sqrt{0.0017} \approx 4\%$；如果遗传度为100%时，患者一级亲属的再发风险上升到9%；如果遗传度在50%时，患者一级亲属的再发风险下降到2%。由此可见，多基因遗传病的再发风险与疾病的遗传度高低有关。

因此，有了群体患病率和遗传度，即可对患者一级亲属患病率做出适当估计（图6-7）。

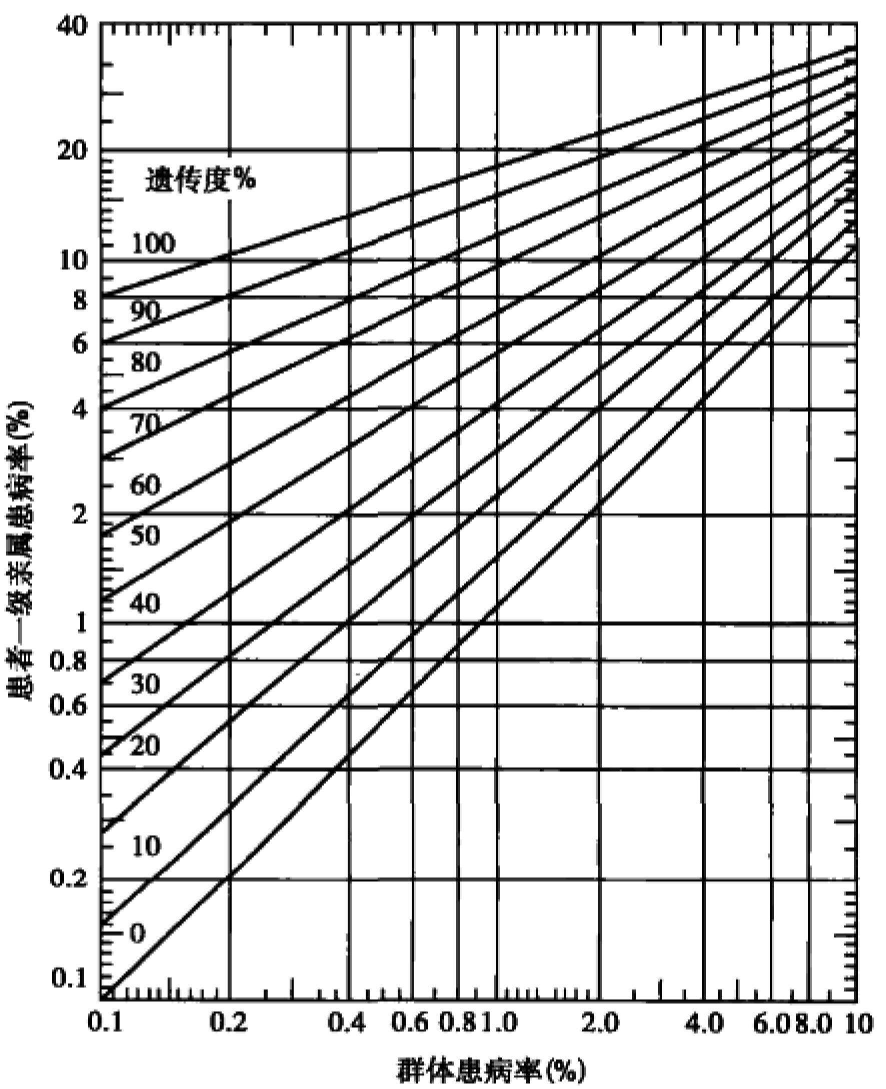


图6-7　群体中患病率、遗传度与患者一级亲属患病率的关系

表6-5 多基因病再发风险估计(Smith 表格)

双亲患者数		0			1			2		
一般群体患病率(%)	遗传度(%)	再发风险率%								
		同胞患者数			同胞患者数			同胞患者数		
		0	1	2	0	1	2	0	1	2
1.0	100	1	7	14	11	24	34	63	65	67
	80	1	8	14	8	18	28	41	47	52
	50	1	4	8	4	9	15	15	21	26
0.1	100	0.1	4	11	5	16	26	62	63	64
	80	0.1	3	10	4	14	23	60	61	62
	50	0.1	1	3	1	3	9	7	11	15

图6-7显示了群体患病率、遗传度和患者一级亲属患病率之间的相互关系,可以从中估计多基因病的发病风险率。例如,无脑畸形和脊柱裂的患病率为0.38%,在图中横轴上查出0.38之点,作一垂直线与纵轴平行,已知此病的遗传度为60%,从图中找出遗传度60%的斜线,把它和0.38的垂直线相交点作一横线在纵轴上的一点近于4,即表明该病的一级亲属患病率接近4%。

有些多基因病,在遗传度相同的情况下,群体患病率不同,发病风险率也不同,同样可以从图中进行估计。

(二) 患者亲属再发风险与亲属中受累人数有关

在多基因遗传病中,当一个家庭中患病人数愈多,则亲属再发风险愈高。例如,一对夫妇表型正常,但头胎出生了一个唇裂患儿以后,再次生育时患唇裂的风险为4%;如果他们又生了第二个唇裂患儿,第三胎生育唇裂风险则上升到10%。说明这一对夫妇带有更多能导致唇裂的致病基因,他们虽然未发病,但他们的易患性更接近发病阈值,因而造成其一级亲属再发风险增高(表6-5)。这一点与单基因病遗传不相同,因为在单基因遗传病中的双亲基因组成已固定,并严格按孟德尔遗传规律遗传,故其后代患病概率不因为已生出几个患者而改变其原有的1/2或1/4发病风险。

(三) 患者亲属再发风险与患者畸形或疾病严重程度有关

多基因遗传病发病的遗传基础是微效基因,故有共显累加效应。在多基因遗传病中如果患者病情严重,证明其易患性远远超过发病阈值而带有更多的易感性基因,与病情较轻的患者相比,其父母所带有的易感基因也多,易患性更接近阈值。因此,再次生育时其后代再发风险也相应增高。例如,一侧唇裂的患者,其同胞的再发风险为2.46%;一侧唇裂并腭裂的患者,其同胞的再发风

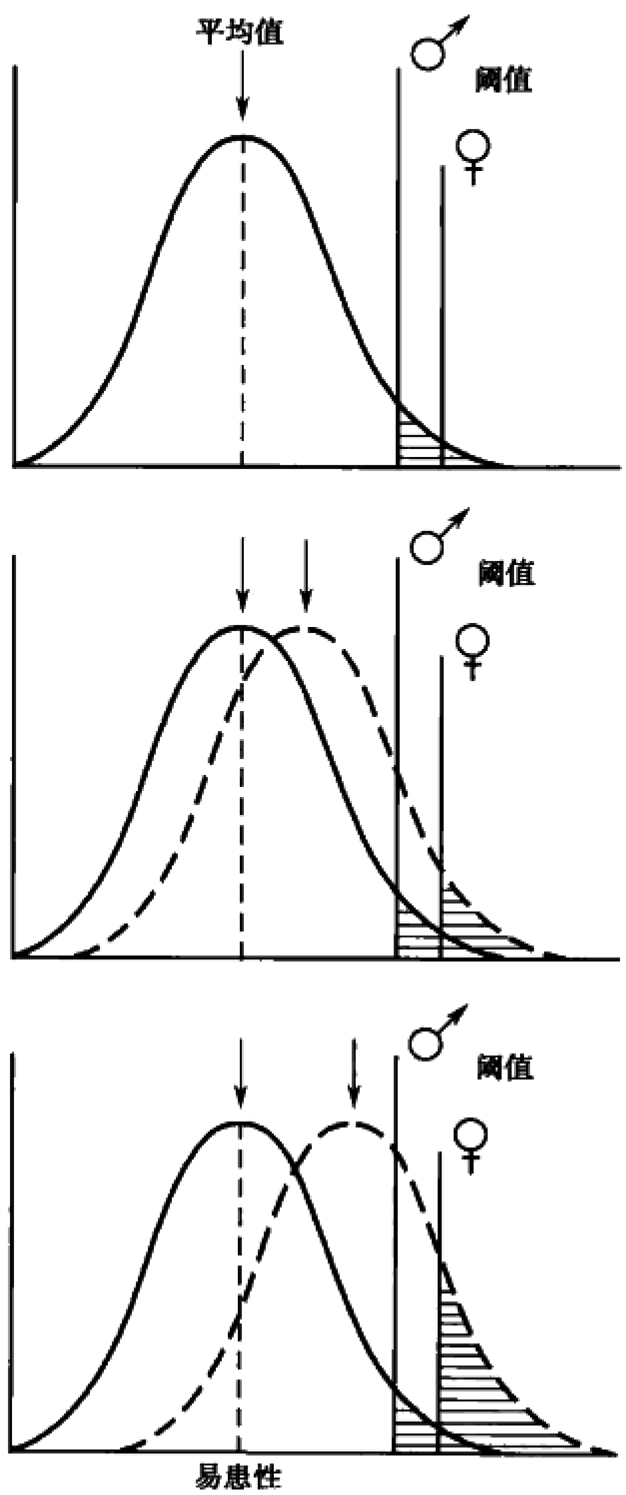


图6-8 群体中先天性幽门狭窄发病阈值有性别差异的易患性分布图

险为4.21%;双侧唇裂加腭裂的患者,其同胞的再发风险为5.74%。这一点也不同于单基因遗传病。在单基因遗传病中,不论病情的轻重如何,一般不影响其再发风险率,仍为1/2或1/4。

(四) 多基因遗传病的群体患病率存在性别差异时,亲属再发风险与性别有关

在某种多基因遗传病的发病上存在性别差异时,表明不同性别的发病阈值是不同的。群体中患病率较低的即阈值较高的性别的先证者,其亲属再发风险相对增高;相反,群体中患病率相对高即阈值较低性别的先证者,其亲属再发风险相对较低。这种情况称为卡特效应(Carter effect)。例如,人群中先天性幽门狭窄男性患病率为0.5%,女性患病率为0.1%,男性比女性患病率高5倍,则男性先证者后代中儿子患病率为5.5%,女儿的患病率是2.4%;而女性先证者后代中儿子患病率高达19.4%,女儿患病率达到7.3%。该结果说明,女性先证者比男性先证者带有更多的易感基因。图6-8显示如果女性患者比例增加时,群体的易患性会增加而右移。

综上所述,在估计多基因遗传病再发风险时,必须全面考虑上述各种情况及有关资料进行具体分析,才能做出切合实际的结论。

第三节　现代医学中多基因遗传病研究的基本思路

单基因疾病虽然种类远较多基因疾病为多,但是由于多基因疾病的发生率大多较高,因此其危害更加严重。如很多心脑血管疾病、老年性痴呆、糖尿病、哮喘等常见病、多发病都有多基因遗传基础。因此对于多基因疾病来说,首先要使这些疾病或性状数量化;其次是利用现代生物学技术(如各种组学技术),筛选出与数量化的多基因疾病的形成与发展相关的基因或多态性遗传标志;并在此基础上,定位和克隆这些基因;再次,需要明确哪些基因是疾病发生的"关键"基因或主基因,哪些基因是不十分"重要"的基因;最后,探讨这些基因在疾病发生、发展中的作用,以及基因与基因之间、基因与环境的相互作用的网络结构,从而揭示多基因疾病的发生机制。

从目前的研究水平来看,多基因疾病患儿的风险预测及产前诊断还相当困难,但随着一些高通量检测技术如基因芯片、蛋白质芯片的应用,易感基因的功能及作用方式将被逐步阐明,可望将多基因疾病的基因诊断用于临床。另外,多基因疾病高风险人群的监测也是一个重要环节,避免使该类人群暴露于危险因素可有效减低发病率。对于一些常见的高发性多基因遗传病,还必须大力提高治疗水平,基因水平的治疗目前还缺乏可靠的理论基础和技术手段,但是随着易感基因的功能及作用方式的阐明,针对性的治疗药物的开发将会达到一个前所未有的高度。

由于绝大多数多基因疾病相当复杂,为了更客观地研究它,许多科学家建立了或试图建立更科学的多基因疾病研究模型,以综合的观点,兼顾环境和遗传两个方面,研究多基因疾病的决定因素。

第四节　几种多基因遗传病

一些常见病或多发畸形,它们的发病率并不低,一般在0.1%~1%,有一定的遗传基础,有家族聚集现象,但患者同胞发病远低于1/2或1/4,一般在1%~10%。

一、精神分裂症

精神分裂症(schizophrenia,SZ,OMIM181500)是人类最常见最令人困惑的疾病之一,全世界的患病率为4%,社会负担居各种疾病的第4位(据WHO统计数据)。家系调查、双生子以及寄养子研究已经证实,SZ属多基因遗传方式(遗传度约为70%~85%)。环境因素如病毒、产伤等也将导致SZ的发病。

(一) 染色体畸变与SZ

有关SZ的染色体畸变与细胞遗传学研究早在20世纪60年代即已开展,已报道的SZ染色体异

常包括:5q11-q13 部分三体、t(2;18)(p11.2;p11.2)易位、t(1;7)(p22;q22)易位、9p11-9q13 倒位、4p15.2-q21.3 倒位、5p14.1 部分三体、8 号三体、8q24 脆性位点、10q24 脆性位点、22q11.1 缺失、5q21-q23.1 缺失以及性染色体非整倍体等。这些染色体畸变只出现在个别 SZ 案例中,并非 SZ 特异性的变异。因而,SZ 与染色体畸变之间没有明显关联。但上述染色体畸变部位的发现,有助于确定 SZ 易感基因。

(二) SZ 候选基因的研究

近年来,应用关联分析方法和全基因组扫描技术,已发现众多基因或位点可能是精神分裂症的易感基因或候选区域。

1. *DRD3* 基因　多巴胺是一种非常重要的神经递质,对调节人体的精神-神经活动具有重要作用。多巴胺过量一直被认为是导致精神分裂症的主要原因,故多巴胺受体基因亦被认为是精神分裂症的重要候选基因。鉴于临床上的许多精神分裂症的治疗药物均为多巴胺 D2 受体(dopamine D2 receptor,DRD2)的阻断剂,*DRD2* 受体基因也因此成为精神分裂症易感基因的候选对象。①*DRD2* 基因位于 11q22.1-22.3,大多数资料显示 *DRD2* 与精神分裂症的易感性无关,但日本患者的 *DRD2* 基因第 141 位 C 碱基缺失频率显著降低,而英国白人患者的第 141 位 C 缺失显著升高;②*DRD3* 基因位于 3q13.3,在第 1 外显子的第 9 个密码子存在由 Ser/Gly 替代形成的 BalⅡ限制性片段长度多态性位点。*DRD3* 主要在端脑、伏隔核、Callegia 岛以及其他边缘系统(如嗅体、海马和乳头体)进行特异表达。与思维、情感等功能有关。研究表明 *DRD3* 的 Ser/Gly BalⅡ限制性片段多态性与精神分裂症的发生存在相关。因此,*DRD3* 基因成为精神分裂症重要的候选基因;③*DRD4* 基因位于 11p15.5,与 *DRD2* 和 *DRD3* 都有明显的同源性,*DRD4* 第 521 位 C→T 多态性以及第 3 外显子 48bp 重要片段的多态性显示与精神分裂症的微弱关联;④迄今为止,尚无证据证明 *DRD1* 和 *DRD5* 两个基因与精神分裂症相关。

2. *5-HTR2A* 基因　神经递质中的另一种重要成分是 5-羟色胺,通过受体介导来调节人的神经活动。在人体中,5-HTR 由多种类型蛋白质组成一个蛋白家族。其中 *5-HTR2A* 基因定位于 13q14,其基因产物是由 471 个氨基酸组成的 G 蛋白偶联受体,特异地分布于带状核、嗅结体、新皮质Ⅰ和Ⅴ层、梨状皮质和嗅前体。研究发现,*5-HTR2A* 第 102 位 T→C 的限制性片段长度多态性位点是理想的遗传标记。目前临床上使用的一些抗精神分裂症新药,均是特异性地作用于 5-HTR2A 而产生药效的,故 *5-HTR2A* 基因可能与精神分裂症的病理变化有关。

3. *HLA* 基因　位于 6p21.3 的 *HLA*(人类白细胞抗原)基因是人类基因组中多态性最丰富的基因群,直接决定免疫排斥反应。某些精神分裂症亚型患者存在自身免疫现象,从而推测 HLA 可能参与精神分裂症的发病过程。大量研究证明,HLA-A1、-A2、-A9、-B5、-CW4、-DR8 等与精神分裂症呈正相关,HLA-DR4、-DQB1 与精神分裂症呈负相关。

4. *KCNN3* 基因　人脑和小鼠脑中存在编码一种较小的钙激活钾离子通道蛋白的 cDNA 家族,分别为 KCNN1、KCNN2 和 KCNN3。人 *KCNN3* 基因(1q21.3)编码 731 个氨基酸长度的多肽链,在基因内靠近 5′端的区域含有两个 CAG 三核苷酸重复序列,第 2 个 CAG 重复序列的多态性最常见。研究发现,精神分裂症患者中较长 CAG 重复等位片段的频率显著高于正常人,*KCNN3* 基因较长的 CAG 重复片段与精神分裂症之间可能存在中等强度的相关。

其余候选的精神分裂症易感基因(候选区)包括:*MTHFR*(1p36)、*RGS4*(1q22)、*CHI3L1*(1q32)、*DISC1*(1q42)、*ERBB4*(2q33)、*SYN2*(3p25)、*PMX2B*(4p13)、*EPNR*(5q33)、*DTNBP1*(6p23)、*NOTCH4*(6p21)、*TRAR4*(6q23)、*NRG1*(8p22-p11)、*GRIN1*(9q34)、*BDNF*(11p13)、*FYXD6*(11q14-q21)、*DAO*(12q24)、*NOS1*(12q24)、*G72*(13q34)、*AKT1*(14q32)、*CHRNA7*(15q14)、*SLC6A4*(17q11-q12)、*GNAL*(18p)、*C3*(19p13)、*APOE*(19q13)、*COMT*(22q11)、*ZDHHC8*(22q11)、*PRODH*(22q11)、*RTN4R*(22q11)以及 CAG/CTG 三核苷酸重复序列。

随着人类精神疾病基因组计划中引用 DNA 芯片的研制及 SNPS 图谱绘制,人们必定能确定精神

分裂症的致病基因，从而提高精神疾病诊断的可靠性，揭示 SZ 发病的分子机制，并为新药研制、个性化治疗提供理论依据。

二、糖尿病

糖尿病（diabetes mellitus，DM）是一种常见多发的与遗传因素有关的复杂性疾病。其患病人数正随着人民生活水平提高、人口老化、生活方式改变以及诊断技术的进步而迅速增加。欧美人群 DM 患病率达 2% ~3%，我国已上升到 3.1%（1996）。据 1997 年 WHO 报告，全世界约有 1.35 亿 DM 患者，预测到 2025 年将上升到 3 亿。在发达国家 DM 是继心血管疾病和肿瘤的第三大非传染性疾病。DM 已成为世界上第五位死亡主因，成为严重威胁人类健康的全球性公共卫生问题。

（一）糖尿病的分类及特征

DM 的病因及发病机制尚未完全阐明。胰岛病变致胰岛素分泌缺乏或延迟，抗胰岛素抗体产生，胰岛素受体缺陷或受体靶组织对胰岛素敏感性降低等构成 DM 发病主要环节。临床表现为以慢性血糖升高为特征的碳水化合物、蛋白质、脂肪代谢紊乱的综合征。1999 年 WHO 咨询委员会采纳了美国糖尿病协会 1997 年提出的关于更新 DM 分型和诊断标准的建设及报告内容。新的 DM 分型中废除 DM 过去沿用的胰岛素依赖性 DM（IDDM）和非胰岛素依赖性 DM（NIDDM）的名称，并以阿拉伯数字 1 和 2 取代过去Ⅰ型和Ⅱ型 DM 中的罗马数字Ⅰ和Ⅱ，取消与营养相关性糖尿病（MRDM），将其归为特殊类型中胰腺外分泌疾病所致的 DM。2 型 DM 占 90% ~95%。

1 型 DM 属于自身免疫性疾病。这类患者由于胰岛 β 细胞膜上 HLA-Ⅱ类基因异常表达，使得 β 细胞成为抗原递呈细胞，在环境因素（病毒感染等）作用下，免疫反应被激活，产生自身抗体，导致胰岛细胞炎症，演变而成为 DM。

2 型 DM 的发病多为自主神经类型，表现为副交感神经张力增加，交感神经张力减弱导致低血糖倾向及多吃、肥胖。2 型 DM 患者随年龄增长，出现胰岛 β 细胞数目减少，胰岛素分泌缺陷或终末器官对胰岛素产生抗性，导致糖尿病。老化过程中胰岛素原合成减少 16% ~39%。

（二）糖尿病发生的遗传因素

DM 的发病具有遗传基础，95% 以上属于多基因遗传病，并且环境因素的影响也很重要，但却出现很强的遗传异质性。例如，1992 年 Ballinger 首次报道了一个 DM 家系，有明显母系遗传特点，伴有感觉神经耳聋，在该家系中发现一例 mtDNA10.4Kb 缺失，相继 Van den Ouwd 和 Reardon 在同样存在母系遗传伴感觉神经耳聋的家系中，又发现了 mtDNA 中 tRNA 基因 3243bpA→G 的突变。而且该种突变已累积报道 45 个家庭 199 例 2 型 DM 患者带有同样突变位点，其中 82% 患有 DM。说明该位点与 2 型 DM 有相关性。在临床中 1 型和 2 型 DM 是完全不同的疾病，其病因、病程和遗传学也有很大差异。

1. 1 型 DM 的发病　筛查 1 型 DM 的易感基因仍然是探索该病病因的主要途径。Davies 等对 1 型 DM 采用患病同胞分析方法，用分布在整个基因组上的 290 个遗传标志进行基因组扫描，在众多筛查的基因型（55 000 个）中，患病同胞对显示显著共有标志者，存在 20 个不同染色体区，其中有 10 个候选区显示共有标志的显著性达 $P<0.005$ 水平，如表 6-6 所示。

经分析可见 HLA 多态性对 $IDDM_1$ 基因座有强烈易感效应。DQA_1*0301、DQB_1*0302、DQB_1*0301 与 1 型 DM 有关联，DQα 链 52 位为非精氨酸、DQβ 链 57 位为非天冬氨酸时显示对 1 型 DM 有强烈易感性。DRB_1 的 DR_3 与 DQB_1*0201 之间有连锁不平衡，DRB_1 的 DR_4 与 DQA_1*0301 之间有连锁不平衡，所以 DR_3 和 DR_4 也显示与 1 型 DM 关联。同时又查出两个可信的新确定的基因座位：$IDDM_4$ 和 $IDDM_5$。

2. 2 型 DM 的发病　2 型 DM 是异质性很强的多基因遗传病，并且受环境因素的重要影响。对 2 型 DM 易感基因的研究目前主要采用两种方法，即候选基因法和以基因组扫描为基础的定位克隆法。国际上已经研究过近 250 多种候选基因，但只发现少数几个在特殊的 DM 类型中呈现出一定的相关

表6-6 经患病同胞对分析1型DM易感基因

基因座位	染色体定位区域	被选作的遗传标记	最高lod值	相对风险	易感性(%)
$IDDM_1$	6p21.3	HLA-DQB	7.3	3.1	42
$IDDM_2$	11P15.5	INSVNTR	2.1	1.3	10
$IDDM_3$	15q26	D15S107	?	?	?
$IDDM_4$、$IDDM_5$	11q13	FGF3	3.4	1.3	?
$IDDM_6$、$IDDM_7$	6q24q27	ESR	2.0	?	?
$IDDM_8$	18q21	D18S64	—	—	—
$IDDM_9$	2q31	D2S152	—	—	—
$IDDM_{10}$	6q27	D6S264	—	—	—
$IDDM_{11}$	3q21-q25	D3S1576	—	—	—
$IDDM_{12}$	10q11.2-q11.2	D10S193	—	—	—
$IDDM_{12}$	14q24.3	D14S48	—	—	—
$IDDM_{14}$	2q33	D2S1391	—	—	—
$IDDM_{13}$	2q34	D2S301	—	—	—
$IDDM_{14}$	未定	未定	—	—	—
$IDDM_{14}$	6q21	D6S238	—	—	—

性。Mahtani在观察分析芬兰地区的2型DM家系中发现，12q的D12S1349存在2型DM易感基因，并命名为NIDDM2。另外，载脂蛋白A2被认为是2型DM的候选基因。近年来研究发现在20号染色体长臂上存在许多重要的2型DM的候选基因，已报道的有肝细胞核因子4a(HNF-4aMODY1)、磷酸烯醇式丙酮酸羧激酶(PCK)基因、agouti signaling protein(ASP)、HNF-3β、磷脂酶C-G1(PLCG1)和CCAAT增强子结合蛋白B(C/EBP-β)等。可以预见，几年内将会出现更多的报道，这些研究必将最终定位和克隆DM的致病基因。

三、哮喘

支气管哮喘(bronchial asthma)或哮喘是一种以气道炎症、气道高反应性和可逆性气道阻塞为特征的呼吸疾病。该病病因复杂，受遗传因素和环境因素影响。患病率和死亡率在很多国家呈上升趋势。世界各地患病率为1%～4%。我国华南地区为0.69%，北京地区为5.29%。以儿童多见，男性发病率略多于女性，农村多于城市，有季节和时间性发病特点，如冬季或夜间常发作。

（一）哮喘的临床特征

哮喘是以过敏原或非过敏原因素引起的支气管反应性增高的疾病。是通过神经体液而导致气道可逆性的痉挛、狭窄。临床上表现为发作时带有哮鸣音的呼气性呼吸困难，可持续数分钟至数小时，可自行或治疗后缓解，严重时可持续数日或数周，反复发作；长期反复发作哮喘可并发慢性支气管炎或肺气肿。

哮喘可大致分为外源性和内源性二大类，但又可根据病因、产生机制和防治不同，分为吸入型(外源性)、感染型(内源性)、运动型、药物型和混合型，以前两型多见。外源性哮喘常于幼年发病，具有明显的对过敏原的多态反应史。内源性哮喘常于成年发病，支气管迷走神经反应性增高，倾向于常年发作，且较为严重。外源性和内源性哮喘在发病过程中可互相影响而混合存在。

各型哮喘共同的支气管病理改变主要是支气管平滑肌痉挛、黏膜水肿及炎细胞浸润，管壁腺体过

度分泌入管腔形成黏液栓，从而引起支气管阻塞。目前对其产生机制尚不清楚。一般认为外源性哮喘产生与免疫反应异常有关，而内源性哮喘多因为自主神经功能紊乱所致。

（二）哮喘发生的遗传因素

近20年来，通过基因组扫描、候选基因技术和连锁分析等方法，确认哮喘是一种复杂的多基因病，有家族聚集性的特点。它的发生、发展由宿主遗传易感性和环境暴露相互作用所决定。研究表明多个染色体上的基因与哮喘相关，主要集中在5q、6p、11q、12q、14q、19q等。

对哮喘的研究表明，过敏性哮喘与特应性（atopy）体质有密切关系。特应性体质者的IgE高亲和力受体基因（$Fc\varepsilon R_1$-B）经与遗传标志D11S917连锁分析，定位于11q13；对一过敏性哮喘的大家系用遗传标志D11S533作连锁分析，lod记分>1，其易感主基因可能定位于11q13-q13.4。位于14q11.12的T细胞受体λ链基因（TCQ2）区中有调节IgE反应的基因，也可能与哮喘有关；另外气道高反应性是哮喘的一个标志，并是强大的危险因素，β_2-肾上腺素受体ADRB2基因第16位和27位发生突变，与哮喘的一些临床特征有关，包括在气道高反应性，1998年Damato在高度同质人群中对ADRB2进行研究发现16位G1y和27位Gln以及ADRB2单倍型与气道高反应性相关，在校正了如特异性和总IgE水平的潜在混杂后，这一相关仍然存在。在我国福建省诏安县渔民的近亲婚配群体中，过敏性哮喘的发病率高达5.6%，这是一个很好的遗传学研究背景。进一步分析可能确认出哮喘致病基因。

表6-7所列是目前研究的一些哮喘候选基因。

表6-7 某些被怀疑的哮喘候选基因

候选基因	染色体定位区域	所研究的种族或民族、地区哮喘患者
KCNS3（电压门控延迟整流钾通道亚家族S成员3）	2p24	533个中国安徽省安庆市家系
HNMT（组氨酸-N-甲基转移酶）	2q22	美国、日本、德国儿童、印度患者
MUC7（唾液黏蛋白7）	4q13-q21	北欧患者
IL13（白介素-13）	5q31	荷兰、英国、日本患者
IL12B（白介素-12B）	5q31-q33	美国患者
SCGB3A2（分泌珠蛋白家族3A成员2）	5q31-q34	日本患者
ADRB2（B_2-肾上腺素受体）	5q32-q34	美国患者
HLA-G（人类白细胞抗原-G）	6p21	美国、荷兰患者
PLA2G7（磷脂酶A2第7组）	6p21.2	德国、英国患者
TNFA（肿瘤坏死因子A）	6p21.3	美国、日本、韩国患者
CCL24（趋化因子CC基序配体24）	7q11.23	韩国患者
UGB（子宫珠蛋白）	11q12.3-q13	澳大利亚患者
HLA-DRB1（人类白细胞抗原-DRB1）	6p21.3	英国患者
STAT6（信号转导及转录活化蛋白6）	12q13	德国、美国、加拿大患者
PHF11（PHD指蛋白11）	13q14	英国患者
IL4R（白介素-4受体）	16p12.1-p11	荷兰患者
ADAM33（解联蛋白与金属蛋白酶结构域33）	20p13	美国患者

2004年4月，芬兰学者Laitinen等发现位于人类第7号染色体短臂（7p）包含2个基因的133kb区段，可能是决定哮喘（asthma）的易感基因。其中1个基因编码孤立的G-蛋白偶联受体，命名为

GPRA(G protein-coupled receptor for asthma susceptibility)。GPRA 蛋白的类型在哮喘患者和健康人的支气管分布存在显著差异。对3个芬兰和加拿大核心家系的关联分析表明,具有 *GPRA* 基因 SNP 单倍型者要么 IgE 升高,要么患有哮喘病。小鼠动物模型显示,其同源 *GPRA* 基因表达上调。因此,*GPRA* 基因可能是引起特应性或哮喘的分子病因。

Summary

Although many human diseases are inherited as simple Mendelian traits or are associated with chromosomal abnormalities, most of the common diseases of adult life(*e.g.*, diabetes mellitus, hypertension, schizophrenia) and most common congenital malformation(*e.g.*, cleft lip, cleft palate, neural tube defects) are not. Yet, there is considerable evidence that genetic factors are important in their etiology. This genetic predisposition to disease is thought to reflect the cumulative effect of genetic variation at several and possibly multiple loci, each with a relatively small effect on phenotype. We define as polygenic those traits or diseases caused by the impact of many different genes, each having only a limited individual impact on phenotype, and as multifactorial those traits resulting from the interplay of multiple environmental factors with multiple genes. In practice, the terms are often used interchangeably. Polygenic traits are usually quantitative rather than qualitative in nature and are frequently distributed continuously in the population, often in a more or less normal frequency distribution, *e.g.*, height and blood pressure.

(左 伋)

参考文献

1. Nussbaum RL, McInnes RR, Willard HF. Genetics in Medicine. 7th ed. Elsevier(Singapore)Pte Ltd, 2007
2. http://www.ncbi.nlm.nih.gov
3. Scriver CR. The metabolic and molecular bases of inherited disease. 8th ed. New York: McGraw-Hill Companies, 2001
4. Thomson G, Esposito MS. The genetics of complex diseases. Trend Cell Biol, 1999, 9(12): 16-20
5. Taylor JG. Using genetic variation to study human disease. Trend Mol Med, 2001, 7: 507-512
6. Rossant J, Mckerlie C. Mouse-based phenogenomics for modelling human disease. Trend Mol Med, 2001, 7: 502-507
7. Cardon LR, Bell JI. Association study designs for complex diseases. Nat Rev Genet, 2001, 2: 91-99
8. Brookes A. Rethinking genetic strategies to study complex diseases. Trend Mol Med, 2001, 7: 512-516
9. Laitinen T, Polvi A, Rydman P, et al. Characterization of a common susceptibility locus for asthma-related traits. Science. 2004; 304(5668): 300-304

第七章 群体遗传

任何一个通过有性生殖繁衍的物种的个体都不可能孤立地存在,总是依附于某个由一群可相互交配的个体所组成的群体。关于群体的遗传学结构、组成及其相关理论的诸多问题是无法仅仅从个体水平的遗传结构来解释的。例如,为什么导致血友病的凝血因子Ⅷ和凝血因子Ⅸ的等位基因在所有种族中都十分罕见,而镰形红细胞贫血症在非洲赤道的某些地区却十分普遍?由于生活环境的改变,美洲黑人后裔的镰形红细胞贫血症的发病率将出现怎样的变化?增加或减少近亲婚配率的后果是什么?等等。所有这些决定群体的遗传组成及其随时间和空间的变化规律性问题都属于群体遗传学的研究范畴。

群体遗传学的目标,就是探索群体的遗传组成以及引起群体遗传组成发生变化的动力。这就需

要调查下述事实：

1. 群体中携带不同基因型个体间的婚配形式。个体间的婚配包括随机婚配(random mating)、近亲婚配(consanguineous mating)和选型婚配(assortative mating)等。

2. 群体间的混合、迁移或分群对群体遗传结构的影响。

3. 突变和遗传重组引起的群体遗传变异的速率。

4. 自然选择对群体遗传结构变化速率的影响。

5. 有限容量的群体中，基因的遗传漂变(genetic drift)对群体遗传结构的影响。

群体遗传学既是一门实验科学，又是一门理论科学。一方面，通过合理地设计实验并观察实验结果，从而能够描述群体连续变异规律，并估计婚配形式、突变以及自然选择的遗传参数。另一方面，依据20世纪30年代RA Fisher，S Wright和JBS Haldane建立的经典群体遗传学理论，可以预测群体的遗传组成及其变化，并研究当不同的因素作用于群体时，群体的遗传结构随时间和空间变化的基本规律。

遗传变异是群体遗传学研究中最基本的研究对象。但通常只能观察到个体的表型变异。有时，观察到的表型变异与遗传变异一致，譬如人类的MN血型，它由一对等位基因L^M和L^N决定，M型和N型是纯合体，MN型是杂合体，对该性状而言，基因型与表型一致，几乎与环境变化无关。但对于身高、体重、血压等大多数性状而言，个体的表型与基因型间存在非常复杂的关系，且与环境的变化紧密相关，目前不可能得到决定这些性状的基因的准确描述。因此，多年来人类群体遗传学的实验研究主要局限于对单基因性状的研究。本章我们主要讨论这类基因型与表现型呈一一对应关系的质量性状在群体中的遗传组成及其变化规律。

第一节　基因及基因型频率在群体中的平衡：Hardy-Weinberg 平衡律

一、Hardy-Weinberg 定律的表述

我们知道，由于等位基因间的显、隐性关系，群体中有变异的个体与正常个体婚配时，显性基因的作用能够把隐性基因的作用掩盖起来。因此，群体遗传学首先面临的一个问题就是：当群体容量充分大，且突变的等位基因频率很低时，因突变导致的隐性变异是否会逐渐消失呢？Hardy(1908年)与Weinberg(1909年)先后独立地证明：如果一个群体满足下列所有条件：

1. 基因型频率没有性别差异。

2. 群体容量无限大。

3. 随机婚配，即群体内所有个体间婚配机会完全均等，每个配子进入合子的机会也完全均等。

4. 没有突变和回复突变，也没有来自其他群体的基因交流。

5. 没有任何形式的自然选择。

那么，该群体常染色体基因座上的基因型比例，经过一个世代的随机交配以后，仍可以维持不变。此即所谓的遗传平衡定律(law of genetic equilibrium)。这一基因或基因型频率不随世代变化而处于平衡状态的现象，被尊为Hardy-Weinberg定律。这一发现阐明了生物群体的一条最重要的遗传学性质：基因的遗传机制即基因传递过程中的随机重组和分离，其本身并不影响群体中保持遗传变异的平衡机制。这一规律的发现奠定了现代群体遗传学最重要的理论基础。根据这条定律可以知道，虽然显性基因的作用可以掩盖隐性基因的作用，但是各基因型的比例不变，所以隐性变异不会因此而逐渐消失。下面我们分别从双等位基因和复等位基因的情形出发，导出该定律。

二、双等位基因的 Hardy-Weinberg 定律的推证

设群体某一基因座上有一对等位基因A、a，三种可能的基因型分别是AA、Aa和aa，f_{AA}、f_{Aa}、f_{aa}分

别为相应基因型的频率。由于纯合子 AA 全部为等位基因 A，而杂合子 Aa 含有一半的等位基因 A，故群体中等位基因 A 的频率 p 为：

$$p = f_{AA} + \frac{1}{2}f_{Aa} \tag{7-1}$$

同理，群体中等位基因 a 的频率 q 为：

$$q = f_{aa} + \frac{1}{2}f_{Aa} \tag{7-2}$$

同时
$$p + q = f_{AA} + f_{Aa} + f_{aa} = 1$$

在一个无选择、无突变、无限大的随机交配群体内，考察某一常染色体基因座上的一对等位基因。假定精子和卵子中等位基因 A 的频率均为 p，等位基因 a 的频率均为 q，见表 7-1，则第二代群体中基因型为 AA 的频率为 $p \times p = p^2$，同理，第二代基因型为 aa 的频率为 $q \times q = q^2$，而第二代杂合子 Aa 的频率为 $p \times q + q \times p = 2pq$，因此，三种基因型频率之比为：$AA : Aa : aa = p^2 : 2pq : q^2$。由式 7-1 和式 7-2 知，下一世代的等位基因 A 的频率为 $p = p^2 + pq = p(p+q) = p$，等位基因 a 的频率为 $q = q^2 + pq = q(p+q) = q$，因此，等位基因频率不随世代改变，且基因型频率也保持 $AA : Aa : aa = p^2 : 2pq : q^2$，不随世代改变。

表 7-1 一对等位基因配子的随机结合，得出三种基因型的频率

		卵子	
		$A(p)$	$a(q)$
精子	$A(p)$	$AA(p^2)$	$Aa(pq)$
	$a(q)$	$Aa(qp)$	$aa(q^2)$

三、复等位基因的 Hardy-Weinberg 遗传平衡式

A_1、A_2、A_3 为群体某一基因座上的复等位基因，$f_{A_1A_1}$、$f_{A_1A_2}$、$f_{A_1A_3}$、$f_{A_2A_2}$、$f_{A_2A_3}$、$f_{A_3A_3}$ 分别为该座位上六种基因型的频率。由于纯合子全部为同一种等位基因，而杂合子含有各一半的相应等位基因，故可得：

等位基因 A_1 的频率 p 为：

$$p = f_{A_1A_1} + \frac{1}{2}f_{A_1A_2} + \frac{1}{2}f_{A_1A_3} \tag{7-3}$$

等位基因 A_2 的频率 q 为：

$$q = f_{A_2A_2} + \frac{1}{2}f_{A_1A_2} + \frac{1}{2}f_{A_2A_3} \tag{7-4}$$

等位基因 A_3 的频率 r 为：

$$r = f_{A_3A_3} + \frac{1}{2}f_{A_1A_3} + \frac{1}{2}f_{A_2A_3} \tag{7-5}$$

仍然考虑一个无选择、无突变、无限大的随机交配群体。则各种基因型频率由表 7-2 可得。

表 7-2 复等位基因配子的随机结合，得出 6 种基因型频率

		卵子		
		$A_1(p)$	$A_2(q)$	$A_3(r)$
精子	$A_1(p)$	$A_1A_1(p^2)$	$A_1A_2(pq)$	$A_1A_3(pr)$
	$A_2(q)$	$A_2A_1(qp)$	$A_2A_2(q^2)$	$A_2A_3(qr)$
	$A_3(r)$	$A_3A_1(rp)$	$A_3A_2(rq)$	$A_3A_3(r^2)$

因此下一世代满足 Hardy-Weinberg 平衡的基因型频率为：

$$(p+q+r)^2=p^2+q^2+r^2+2pq+2pr+2qr=1$$

由于

$$1=p+q+r=(p+q+r)^2=\sum_{i=1}^{3}f_{A_iA_i}+\sum_{i<j}f_{A_iA_j} \tag{7-6}$$

故一旦达到 Hardy-Weinberg 平衡后,等位基因频率和基因型频率均不再随世代发生改变。

一般地,对于具有更多等位基因的基因座,我们有:

设群体中某一基因座上有复等位基因分别为 A_1、$A_2\cdots A_m$,$p_i(i=1,\cdots,m)$表示等位基因 A_i 的频率,$f_{A_iA_i}(i=1,2,\cdots,m)$、$f_{A_iA_j}(i<j,i=1,2,\cdots,m,j=2,3,\cdots,m)$分别为该座位上所有可能的纯合和杂合基因型的频率,则有:

等位基因 $A_i(i=1,2,\cdots,m)$的频率为:

$$p_i=f_{A_iA_i}+\frac{1}{2}\sum_{j=1,i\neq j}^{m}f_{A_iA_j}(i=1,2\cdots,m) \tag{7-7}$$

且满足 Hardy-Weinberg 平衡的基因频率和基因型频率的关系为:

$$1=\sum_{i=1}^{m}(p_i)=\left(\sum_{i=1}^{m}p_i\right)^2=\sum_{i=1}^{m}p_i^2+2\sum_{i<j}^{m}p_ip_j \tag{7-8}$$

式中,$\sum_{i=1}^{m}p_i^2$ 称为群体中基因的纯合度,而 $2\sum_{i<j}^{m}p_ip_j=1-\sum_{i=1}^{m}p_i^2$ 称为群体中基因的杂合度。

杂合度是群体在某一基因座位上遗传变异程度的一个测度,由该座位所有杂合子在群体中所占的频率来表示。如果某个等位基因具有很高的频率而其他等位基因的频率都接近于零,则群体的杂合度将很低,因为绝大多数的个体将是同一等位基因的纯合体。如果同一座位的所有等位基因都具有相同的基因频率,则该座位上群体的杂合度达到最大。

第二节 Hardy-Weinberg 平衡律的应用

一、估计基因频率和杂合度

(一) 常染色体基因频率和杂合度的估计

1. 共显性等位基因的基因频率估计　上海居民中,调查 1788 人的 M-N 血型,其中 397 人是 M 型,861 人是 MN 型,530 人是 N 型。根据 M-N 血型的遗传模式可知,每个 M 型个体带有两个 L^M 基因,每个 MN 型个体带有一个 L^M 基因和一个 L^N 基因,每个 N 型个体带有两个 L^N 基因。就 L^M 和 L^N 这对等位基因而言,1788 人共有 3576 个基因,所以

L^M 基因频率的估计值为:$p=\dfrac{397\times2+861}{3576}=0.4628$

L^N 基因频率的估计值为:$q=\dfrac{530\times2+861}{3576}=0.5372$

2. 显、隐性等位基因的基因频率和杂合度的估计　具有显、隐性关系的等位基因,按 Hardy-Weinberg 律,其隐性基因频率可由隐性纯合基因型频率经开方后求得。例如尿黑酸尿症为常染色体隐性遗传病,约 1 000 000 个儿童中有 1 个患儿,其发病率 $x=0.000001$。由于尿黑酸尿症为隐性纯合体(aa)致病,其双隐性基因型频率为隐性基因频率的平方,即 $q^2=x$。故

隐性基因频率为:$q=\sqrt{x}=\sqrt{0.000001}=0.001$

正常显性基因频率为:$p=1-q=0.999\approx1$

杂合子频率为：$2pq \approx 2 \times 1 \times 0.001 = 0.002$

因此，当隐性基因频率很低时，p 接近于 1，$2pq$ 近似于 $2q$，此时杂合子（即携带者）的频率是隐性基因频率的 2 倍，即 $2q$，比患者频率（q^2）高得多。杂合子频率与纯合子频率之比为 $2pq/q^2 = 2/q$，即杂合子频率是隐性纯合子患者频率的 $2/q$ 倍。这意味着隐性基因频率愈小时，杂合子频率相对于隐性纯合子频率的倍数愈高。上例尿黑酸尿症，人群中尿黑酸尿症携带者频率是患者频率的 $2/0.001 = 2000$ 倍。

（二）X 连锁基因频率的估计

X 染色体上的基因频率如按女性群体数据计算，其方法同常染色体基因频率的计算；如按男性群体数据计算，因男性是半合子，所以男性群体的表型频率即为此基因频率。从男性和女性群体数据分别得出的基因频率基本上是相同的。

试以 Xg 血型为例，这是由 X 染色体上一对等位基因 Xg^a 和 Xg 决定的，其中 Xg^a 为显性。用血清学方法可检出有 Xg^a 基因的个体对抗 Xg^a 血清呈阳性反应，表现为 $Xg^{(a+)}$，只具有 Xg 基因的个体则呈阴性反应，表现为 $Xg^{(a-)}$。对 589 个白人进行调查发现：1）在 298 个男性中有 188 人为 $Xg^{(a+)}$，占 63.1%；有 110 人为 $Xg^{(a-)}$，占 36.9%。男性表现型百分率就是他们的基因型 Xg^aY 和 XgY 的百分率，同时也是 Xg^a 基因和 Xg 基因的频率，故 Xg^a 的基因频率为 0.631。2）在 291 个女性中有 260 人对抗 Xg^a 血清呈阳性反应，占 89.3%；有 31 人对抗 Xg^a 血清呈阴性反应，占 10.7%。因女性 $Xg^{(a+)}$ 个体中包含有两种基因型 Xg^aXg^a 和 Xg^aXg，在 Xg^aXg 基因型中还包含有隐性 Xg 基因，所以要计算女性中 Xg 血型的基因频率应按照估计常染色体基因频率的方法。Xg 的基因频率 $q = \sqrt{0.107} = 0.327$，故 Xg^a 的基因频率为 $1 - 0.327 = 0.673$，与男性中的频率相近。

（三）复等位基因的基因频率估计

1. 共显性复等位基因的基因频率估计　对于具有完全共显性复等位基因的基因座位而言，各种基因型可以精确确定，则可按照式 7-7 求出各等位基因频率，不再详述。

2. ABO 血型复等位基因的基因频率的估计　人类的 ABO 血型系统，是共显性与显性系统的混合体。其等位基因的确定要复杂些。通常我们可以获得如表 7-3 的样本：

表 7-3　人类 ABO 血型系统样本

表现型	A	AB	B	O
基因型	AA/AO	AB	BB/BO	OO
观察值	N_A	N_{AB}	N_B	N_O

这里 N_A 是 N_{AA}（基因型为 AA 的个体数）与 N_{AO}（基因型为 AO 的个体数）之和，N_B 是 N_{BB}（基因型为 BB 的个体数）与 N_{BO}（基因型为 BO 的个体数）之和。如果三种等位基因 A、B、O 的基因频率 p_A、p_B、p_O 已知，在满足 Hardy-Weinberg 平衡的条件下，表现型为 A 型，基因型为 AA、AO 的个体数分别为：

$$N_{AA} = N_A \times \frac{p_A^2}{p_A^2 + 2p_Ap_O}$$

$$N_{AO} = N_A \times \frac{2p_Ap_O}{p_A^2 + 2p_Ap_O} \tag{7-9}$$

同理可得表现型为 B 型、基因型为 BB、BO 的个体数分别为：

$$N_{BB} = N_B \times \frac{p_B^2}{p_B^2 + 2p_Bp_O}$$

$$N_{BO} = N_B \times \frac{2p_Bp_O}{p_B^2 + 2p_Bp_O} \tag{7-10}$$

于是，按照式 7-3 ~7-5，可以得到 A、B、O 三种等位基因的频率为：

$$p_A = \frac{2N_{AA} + N_{AO} + N_{AB}}{2N}$$

$$p_B = \frac{2N_{BB} + N_{BO} + N_{AB}}{2N}$$

$$p_O = \frac{2N_{OO} + N_{AO} + N_{BO}}{2N} \tag{7-11}$$

式中 N 为总样本容量。由于 A 对 O、B 对 O 均为显性效应，上式中只有 N_{AB} 和 N_{OO} 是确定的，我们无法直接应用上式求出等位基因 A，B，O 的基因频率。运用期望极大法（The Expectation Maximization Algorithm，EM 算法）的迭代法，可以求得所需的三种等位基因频率。其做法是，首先任意给出三种基因频率的"初值"，譬如，$p_A = 0.33$，$p_B = 0.33$，$p_O = 0.34$；将这些给定的等位基因频率的初值，连同表 7-3给出的各表现型的个体数，代入式 7-9 和式 7-10，算出各种基因型个体数，再回代入式 7-11，计算出新的等位基因频率；如此反复，直到前后两次计算所得的基因频率差别小于我们事先给定的误差限，迭代停止。这也是极大似然估计的基本思路。

假定我们得到一组观测值如下：

表现型	A	AB	B	O	合计
观察值	862	131	365	702	2060

在满足 Hardy-Weinberg 平衡的条件下，运用上述 EM 算法，经迭代求出：$p_A = 0.281$，$p_B = 0.129$，$p_O = 0.590$。

二、遗传假设的 χ^2 检验方法

（一）χ^2 检验的基本原理

医学遗传学研究中，经常需要将一组观察获得的数据与一组基于某种假定条件（或称原假设）下获得的预期数相比较，根据这两组数据间的差异是否显著，来判断原假设是否正确。如果预期值与观察值十分接近，我们有理由相信原假设是正确的。而如果预期值与观察值相距甚远，我们也有把握拒绝原假设。然而，实践中我们不可避免地会遇到这类情况，即预期值与观察值间的差异既不是非常接近，也不是截然不同，这时 χ^2 检验将有助于我们决定是接受还是拒绝一个原假设。

χ^2 检验的基本原理是，当原假设成立时，χ^2 检验告诉我们这样一个事实：由于随机因素的影响，观察到的数据与理论期望值之间的误差出现的概率能有多大？然而，在原假设成立时，出现任何观察值的概率均不为零，那怎样才能对一个原假设进行取舍呢？通常把出现观察值与预期值之间的误差限的概率小于 5% 作为拒绝原假设的一个标准。

5% 的意思是：我们有 5% 的可能拒绝一个原本是正确的假设。这是我们可以接受的犯第一类错误（拒真）的概率。倘若我们不愿付出这样的代价，即将拒绝原假设的标准定得很低（譬如说 0.1%）。这样，虽说几乎每一个原本正确的假设都会被我们接受（接受概率为 99.9%），但由于原假设的接受域非常宽，将不可避免地导致原本不正确的假设也被我们错误地接受，即犯第二类错误（存伪）的概率将明显增加。因此，一般情况下，将拒绝原假设的标准定得过低或过高都不合适。

下面我们通过实例来了解如何运用 χ^2 检验判断一个原假设的真伪。其中 p 值根据求得的 χ^2 值经查 χ^2 分布表得出。它表示当原假设成立时，重复实验的观察值与预期值之间的误差，出现比当今观察值与预期值之间的误差还要大的概率。通常 $p > 0.05$ 就认为可以接受原假设，即认为观察值与

预期值之间无显著差异。

要正确进行χ^2检验，必须给出正确的χ^2检验自由度(degree of freedom，d. f.)。通常通过下式求得χ^2检验的自由度：

$$d.f. = (分组数 - 1) - (运用观测数据估计的独立参数个数)$$

(二) 利用χ^2检验 M-N 血型基因型的 Hardy-Weinberg 平衡

假设上海人群的 MN 血型分布符合 Hardy-Weinberg 平衡。把前面计算得到的基因频率代入基因型的平衡频率，再乘以总人数，可得到满足 Hardy-Weinberg 平衡的基因型期望数，再将之与各基因型实际人数比较，进行χ^2检验，见表 7-4。

表 7-4 MN 血型资料与遗传平衡间好适度的χ^2检验[4]

	L^ML^M	L^ML^N	L^NL^N	合计
实得数(O)	397	861	530	1788
预期值	np^2	$2npq$	nq^2	n
预期数(C)	382.96	889.05	515.99	1788
$\frac{(O-C)^2}{C}$	0.51	0.88	0.38	$\chi^2_{[1]}=1.77$

$0.1 < p < 0.2$

计算结果，将 0.51、0.88、0.38 相加得$\chi^2_{[1]} = 1.77$，$0.1 < p < 0.2$ 表明在满足 Hardy-Weinberg 平衡条件下，出现表 7-4 这样的偏差的概率大于 0.1，故可以接受三个基因型频率符合 Hardy-Weinberg 遗传平衡的假设。该处χ^2检验的自由度为 d. f. =1，这是因为在计算预期频率时要应用一个基因频率值，而这是根据实际观测值估计出来的，因此χ^2自由度又减去一个，故最终的自由度为(3 - 1) - 1 = 1。

(三) 检验 X 连锁基因频率分别在男、女群体中无差异的假设

由前述可知，通过不同方法计算得出的男女群体 Xg^a 基因频率相近。现通过χ^2检验进一步加以验证。期望值 C_1 和期望值 C_2 是分别按男性群体和女性群体中所得的基因频率得出的期望人数。对此进行 Xg^a 基因频率在男、女群体中有无差异的χ^2检验。结果见表 7-5。

表 7-5 男女群体中 Xg 血型基因频率的期望值

	男性		女性		合计
	$Xg^{(a+)}$	$Xg^{(a-)}$	$Xg^{(a+)}$	$Xg^{(a-)}$	
观察值(O)	188	110	260	31	589
期望值 C_1*	$298\times0.631=188.04$	$298\times0.369=109.96$	$291\times(1-0.136)=251.37$	$291\times(0.369)^2=39.62$	588.99
期望值 C_2**	$298\times0.673=200.55$	$298\times0.327=97.45$	$291\times(1-0.107)=259.86$	$291\times(0.327)^2=31.14$	589
$\frac{(O-C_1)^2}{C_1}$	0.00	0.00	0.29	1.87	2.16
$\frac{(O-C_2)^2}{C_2}$	0.79	1.62	0.00	0.00	2.41

注：* 期望值 C_1：用男性群体的基因频率计算期望人数；** 期望值 C_2：用女性群体的基因频率计算期望人数

观察值与期望值 C_1 的误差的χ^2计算值为 2.16，p 值的范围为：$0.1<p<0.2$；观察值与期望值 C_2 的误差的χ^2计算值为 2.41，p 值的范围为：$0.1<p<0.2$。自由度均为(2 - 1) = 1。

经χ^2检验知，在X伴性遗传中从男性群体中得到的基因频率和从女性群体中得到的基因频率基本上一致，二者并无显著差异。

（四）检验苯硫脲尝味能力为一双等位基因的隐性遗传假设

苯硫脲（phenylthiocarbamide，PTC）的尝味能力有显著的个体差异。大多数人对低浓度PTC溶液就感到苦味（尝味者）；但也有一部分人要在浓度高得多时才感到苦味（非尝味者）。调查白人家庭父母和子女对PTC的尝味能力：父母均为尝味者的家庭共425个，子女中有929人为尝味者，130人为非尝味者，非尝味者表型频率为0.123；父母一方为尝味者另一方为非尝味者的家庭289个，子女中有483人为尝味者，278人为非尝味者，非尝味者表型频率为0.365。

对PTC尝味能力的遗传提出了一种假设，认为它们受控于一个双等位基因座，*TT*和*Tt*为尝味者，*tt*为非尝味者。现在要问这一遗传假设能否解释上述表现型数据。

按上述遗传假设，在尝味者×尝味者这一婚配类型中，由于尝味者的基因型可为*TT*或*Tt*，因此这种表现型婚配类型包含三种基因型婚配类型，即*TT*×*TT*，*TT*×*Tt*和*Tt*×*Tt*。设*T*基因频率为p，t基因频率为q，$p+q=1$。在随机婚配人群中，这三种基因型婚配的机会，在*TT*×*TT*是$p^2 \cdot p^2 = p^4$，在*TT*×*Tt*是$2 \cdot p^2 \cdot 2pq = 4p^3q$，在*Tt*×*Tt*是$2pq \cdot 2pq = 4p^2q^2$。*TT*×*TT*和*TT*×*Tt*将只产生尝味者的孩子。*Tt*×*Tt*将产生尝味者和非尝味者的孩子，其比例为3∶1。在尝味者×非尝味者婚配类型中有两种基因型婚配类型，*TT*×*tt*和*Tt*×*tt*。*TT*×*tt*的婚配频率是$2 \cdot p^2 \cdot q^2 = 2p^2q^2$，他们的子代将都是尝味者。*Tt*×*tt*的婚配率是$2 \cdot 2pq \cdot q^2 = 4pq^3$，他们的子代将为尝味者和非尝味者各半（见表7-6）。

表7-6　双等位基因假设下，给定婚配类型下的尝味者和非尝味者频率

婚配类型	婚配频率	子代	
		尝味者频率	非尝味者频率
TT×*TT*	p^4	p^4	-
TT×*Tt*	$4p^3q$	$4p^3q$	-
Tt×*Tt*	$4p^2q^2$	$3p^2q^2$	p^2q^2
TT×*tt*	$2p^2q^2$	$2p^2q^2$	-
Tt×*tt*	$4pq^3$	$2pq^3$	$2pq^3$

尝味者×尝味者婚配类型中，子代中非尝味者所占频率是：

$$\frac{p^2q^2}{p^4+4p^3q+4p^2q^2} = \left(\frac{q}{p+2q}\right)^2 = \left(\frac{q}{1+q}\right)^2 \tag{7-12}$$

尝味者×非尝味者婚配类型中，子代中非尝味者的频率是：

$$\frac{2pq^3}{2p^2q^2+4pq^3} = \frac{q}{p+2q} = \frac{q}{1+q} \tag{7-13}$$

已知在随机群体样本3643人中，尝味者是2557人，其表现型频率为0.702，非尝味者是1086人，其表现型频率为0.298。按遗传假设，非尝味者为隐性纯合子，因此$q=\sqrt{0.298}=0.546$。由此可得子代中非尝味者的期望频率，在尝味者×尝味者为$\left(\frac{q}{1+q}\right)^2 = \left(\frac{0.546}{1+0.546}\right)^2 = 0.125$。在尝味者×非尝味者为$\frac{q}{1+q} = \frac{0.546}{1+0.546} = 0.353$。从这两个期望频率可分别得到期望值，应用$\chi^2$显著性检验判断期望值与观察值是否有显著差异。

在尝味者×尝味者婚配类型中，子代非尝味者期望值为$(929+130)\times 0.125 = 132.375$，尝味者的期望值为$(929+130)(1-0.125)=926.625$。

$$\chi^2 = \frac{|929 - 926.625|^2}{926.625} + \frac{|130 - 132.375|^2}{132.375} = 0.0487$$

自由度 =1，$0.90 > P > 0.80$。

在尝味者×非尝味者婚配类型中，子代非尝味者期望值为(438 +278) ×0.353 =268.633，尝味者的期望值为(483 +278)(1 -0.353) =492.367。

$$\chi^2 = \frac{|483 - 492.367|^2}{492.367} + \frac{|278 - 268.633|^2}{268.633} = 0.5048$$

自由度 =1，$0.50 > P > 0.30$。

检验结果：P 值远大于 0.05，表明期望值与观察值无显著差异。根据 Hardy-Weinberg 平衡，可以认为苯硫脲尝味能力受控于一个双等位基因座的隐性遗传假设成立。

第三节 影响群体基因频率的因素(一)：突变和选择

前已述及，Hardy-Weinberg 遗传平衡定律仅适用于无选择、无突变、随机婚配、无限大的理想群体。但是，自然界中不可能有无限大的群体，也很难想象群体中存在永远不会发生突变且绝对不受自然选择影响的基因；同时，人类和许多其他物种的自然群体也不可能是真正意义上的随机婚配群体。在这种情况下，群体的基因频率会发生改变，而群体的基因库也会因不同进化因子相互作用而发生进化。因此，严格来讲，自然界中只能有近似满足 Hardy-Weinberg 遗传平衡条件的群体。为此，我们可以从理想群体出发，设法将限制性的适用条件逐个取消，使理论分析逐渐接近于客观的真实群体的情况，最终获得真实群体的遗传结构及其变化的一般规律。本节我们将讨论在突变和自然选择的作用下，群体中的基因频率是怎样改变的。至于影响基因频率变化的其他因素(如基因频率的遗传漂变、群体迁移和混合行为以及群体的分化等)对进化的影响，将留待第四节讨论。

一、突变对群体遗传平衡的影响

对一个给定的群体，导致群体遗传组成发生变异的原因主要有三个方面：①基因突变；②基因间的重组；③基因在群体间的交流。突变使基因的 DNA 发生了可遗传的改变，是一种潜在的能改变群体中等位基因频率的进化力量，也是群体内发生变异的根源。基因的突变率定义为，基因的一种等位形式在某一世代突变成另外等位形式的概率。突变率通常都很低，许多物种中，不同座位间的突变率是不同的。一些环境因素，如化学物质、放射线、传染介质等均有可能增加突变率。假定初始群体的某一基因座全部是等位基因 A 的纯合体，再假定每一世代由等位基因 A 突变成等位基因 a(称为正向突变，forward mutation)的突变率为 1.0×10^{-5}，则在此后第一世代，等位基因 a 的频率将为 1.0×10^{-5}，而等位基因 A 的频率将为 $1.0 \sim 1.0 \times 10^{-5}$；而第二世代，等位基因 a 的频率将为 $2.0 \times 10^{-5} \sim 1.0 \times 10^{-10}$，而等位基因 A 的频率将为 $(1.0 \sim 2.0) \times 10^{-5} + 1.0 \times 10^{-10}$。显然，新的等位基因 a 的频率增长极其缓慢，并且随着可供突变的原等位基因 A 的频率的降低，突变等位基因 a 的增长速率也逐渐减小。

设 μ 为等位基因 A 突变为其他等位基因 a 的突变率，p_0 为 0 世代 A 的基因频率，p_t 为 t 世代 A 的基因频率。不考虑其他改变群体基因频率的因素，则基因频率的变化率为：

$$\frac{dp}{dt} = -\mu p \tag{7-14}$$

对上式经 0-t 世代积分，得：

$$p_t = p_0 e^{-\mu t} \tag{7-15}$$

上式表明，仅考虑突变的影响时，等位基因 A 的频率随世代数的增加呈指数规律下降。图 7-1 给出当 $\mu = 1.0 \times 10^{-5}$ 时，等位基因 A 的基因频率随世代数的变化。

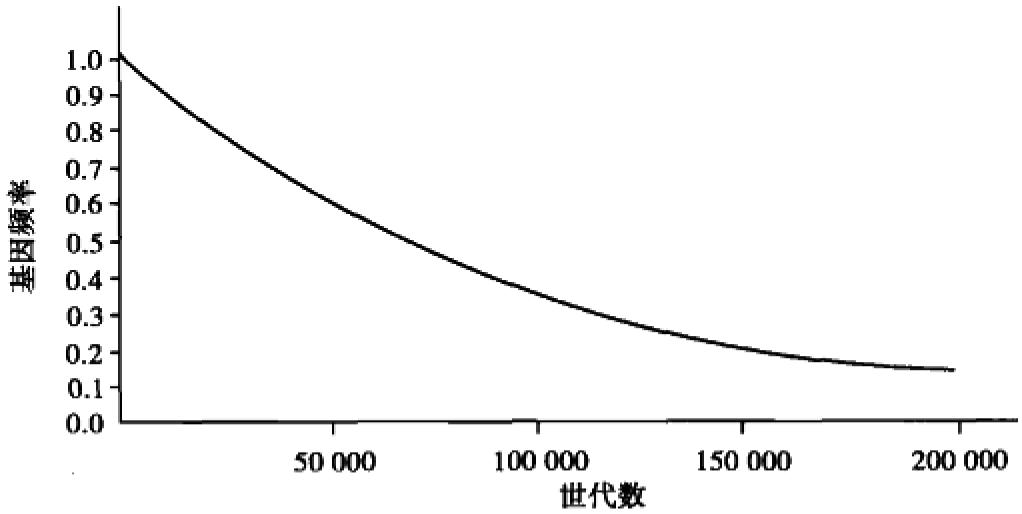


图 7-1　基因突变导致的群体基因频率随世代的变化，突变率 $\mu = 1.0 \times 10^{-5}$

由图 7-1 可知，在突变率很高，且初始群体中仅有等位基因 $A(p_0 = 1.0)$ 的条件下，约 1 万代以后 A 的频率降为 0.9；约 7 万代后 A 的频率降为 0.5，到 200 万代后，A 的频率仍有约 0.15。这表明，由于自发突变的频率通常极低，因此纯粹由突变引起的群体基因频率的改变非常缓慢。

假如将突变视为增加群体新的等位基因种类和频率，而不单纯是降低原有等位基因频率的过程，则这一进化过程将更为缓慢。绝大多数突变率的确定，依赖于等位基因 A 突变成可测定表型变化的等位基因 a。按突变率的严格定义，等位基因 A 的某个特定位置的核苷酸发生突变的频率，往往比实际测得该基因的突变率至少小两个数量级以上。所以尽管有许多突变类型在表型上与传统意义的回复突变（reverse mutation）极类似，但真正意义上的回复突变（即某一特定位置的核苷酸突变成另一核苷酸，然后再经突变回复至原先的核苷酸）几乎是不可能的，或者说概率是极低的。

基因突变是一种可遗传的变异，新的等位基因只有当突变发生时才能产生。因此，突变提供了进化作用的原材料，是进化过程中重要的因素。大部分导致功能改变的基因突变将使个体受到损伤，因而是有害的。而绝大部分的基因突变仅导致基因分子结构的变异，而无功能效应，因而是中性的，个别突变甚至对个体有利，且可能在群体中传播开来。当然，一个突变究竟是有害还是有利，取决于基因与环境间的相互作用。若环境发生了改变，先前有害的突变可以变成有利的突变。

由于突变提供了可遗传的变异材料，为自然选择发挥作用提供了广阔的空间。

二、选择对群体遗传平衡的影响

（一）自然选择的基本概念

Hardy-Weinberg 平衡的重要假定之一，就是群体中所有的个体具有相同的将基因遗传给后代的能力。事实上，构成一个群体的个体之间，在存活能力和生育能力上存在一定差异。达尔文（Charles Darwin）和华莱士（Russel Wallace）在十九世纪中期各自独立地建立了“自然选择”这个极为重要的概念。相比之下，达尔文提出的进化论更为深入，他的卓越贡献使我们深刻认识到了自然选择在物种进化中所起的决定性作用。而《物种起源》一书的出版，也使进化论的思想得到了广泛普及。迄今已由大量观察资料有力地支持了这一学说。

自然选择（natural selection）的含义是由于某种原因，携带某些等位基因的个体比不带这些等位基因的个体具有更多的后代，导致这些基因在下一世代中的频率上升。这样，通过自然选择，携有对生存和繁殖有利的性状的个体得以逐代增加。生物通过这种方式，不断地改变群体中的等位基因和基因型频率，以适应它们赖以生存的环境。

1. 适合度　基因制约着生物体的生理特性或形态结构，而这些生理特性或形态特征又都或多或

少地影响着个体的生活力和繁殖力。因此,绝大多数基因都受到自然选择的作用。在自然选择的过程中,一定的生存能力是必要的,但重要的还是贡献给后代的相对基因数目。带有某一等位基因或基因型的个体可能多留下些子裔,而其他个体可能少留下些子裔。这样,群体中某一基因或某一基因型的频率就会增加,相对地另一基因或另一基因型的频率就会减少。从这个意义上说,选择作用只有发生在育龄期前或育龄期,才会影响群体的基因频率或基因型频率。发生在育龄期之后的选择作用对基因频率或基因型频率的影响将微不足道。

为了对自然选择进行定量研究,将基因型在某一特定环境条件下表现的平均生活力和繁殖力定义为相应基因型的适应值(adaptive value),或适合度(fitness),用 W 表示。适合度指某一基因型与其他基因型相比时能够存活并留下子裔的相对能力。一般将正常的纯合个体的适合度定为1,其他基因型的适合度则用相对生育率(relative fertility)来表示。

例如,根据丹麦的一项调查,软骨发育不全的侏儒108人,共生育了27个儿童。他们的457个正常同胞,共生育582个儿童,故侏儒的相对生育率可表示为:

$$W = \frac{27/108}{582/457} = 0.196 \tag{7-16}$$

相对生育率可以代表适合度。用类似方法可以求得其他各种疾病患者的适合度,如表7-7所示。

表7-7 几种遗传病患者相对适合度的估计(以正常的纯合个体适合度为1.00)

性 状	相对适合度
视网膜母细胞瘤(杂合子)	0
幼儿型黑蒙性痴呆(纯合子)	0
软骨发育不全(杂合子)	0.20
血友病(男性)	0.29
神经纤维瘤病(杂合子)	男0.41;女0.75
慢性进行性舞蹈病(杂合子)	男0.82;女1.25
镰形细胞性状(杂合子)	1.26(在疟疾区)

表中一个有趣的例子是镰形红细胞性状,由于其杂合子对恶性疟原虫的抵抗力较强,因此在非洲的某些地区,杂合子的相对生育率反而较正常人高;而在美洲的黑人人群中却观察不到该种现象。因此,适合度是基因或基因型与环境共同作用的产物。某个基因型在一种环境条件下可能表现得非常适应,在另一种环境条件下可能表现得较为适应或较不适应。可见,环境如果发生改变,基因或基因型的适合度也将随之发生改变。

2. 选择系数 自然选择能产生许多不同的效应:既可以降低甚至剔除有害的遗传变异,也可以保持群体的优势突变;既可以改变基因频率,也可以阻止基因频率的改变;既可以产生和维持群体中的遗传多态性,也可维持遗传的一致性。这些作用取决于基因型的相对适应性和群体中等位基因的频率。选择系数(selective coefficient)指在选择作用下适合度降低的程度,用 s 表示。s 反映了某一基因型在群体中不利于生存的程度,所以 $s = 1 - W$。

选择的结果总是倾向于增加那些具有较高适合度的等位基因和基因型的频率。下面我们讨论不同情况下选择的作用。

(二) 选择对显性基因的作用

1. 基因频率的改变 设显性基因 A 的频率为 p,隐性基因 a 的频率为 q,选择对显性个体不利,选择系数为 s,则经一代选择后,显性基因 A 的频率降低(见表7-8)。

表7-8 显性完全，选择对显性个体不利时基因频率 p 的改变

	基因型			合计
	AA	Aa	aa	
原来频率	p^2	$2pq$	q^2	1
适合度	$W_{11}=1-s$	$W_{12}=1-s$	$W_{22}=1$	
选择后	$p^2(1-s)$	$2pq(1-s)$	q^2	$1-sp(2-p)$
相对频率	$\frac{p^2(1-s)}{1-sp(2-p)}$	$\frac{2pq(1-s)}{1-sp(2-p)}$	$\frac{q^2}{1-sp(2-p)}$	1
基因 A 频率原来为 p		选择后为 $\frac{p-sp}{1-sp(2-p)}$		
基因 A 频率的改变		$\Delta p=\frac{p-sp}{1-sp(2-p)}-p=-\frac{sp(1-p)^2}{1-sp(2-p)}\approx -sp(1-p)^2$		

要注意：经过一个世代的选择作用之后，这三种基因型频率的贡献相加并不等于1。它们的总和 $p^2W_{11}+2pqW_{12}+q^2W_{22}=1-sp(2-p)$，称为该群体的平均适合度。

每一代显性基因 A 频率的改变为：

$$\Delta p=\frac{-sp(1-p)^2}{1-sp(2-p)}\approx -sp(1-p)^2 \tag{7-17}$$

人类自然群体中，显性有害基因的频率很低，$1-sp(2-p)$ 接近于1，所以 $\Delta p-sp(1-p)^2$。由于携带显性基因的个体（纯合或杂合）都要受到选择的作用，故选择对显性基因的作用比较有效。如果没有新的突变产生，显性有害基因较容易从群体中消失。

2. 基因突变率的确定　显性遗传病患者通常都是杂合体，其基因频率为 $2pq$。由于致病基因频率很低，正常等位基因频率 q 接近于1，因此杂合体患者的频率可视为 $2p$。在一个平衡群体中，被淘汰的致病基因将以突变来补偿，假定突变率为 $\nu(a\rightarrow A)$，由于突变产生新的A基因为 νq，同时由于选择作用A基因减少 $\Delta p\approx spq$，故平衡时，$\nu\doteq sp$。当显性遗传疾病致死时，$s=1,\nu=p$，即突变率为发病率的一半。

例如，据在丹麦哥本哈根市的一项调查，几年来在医院所生的94 075个孩子中，有10个患软骨发育不全性侏儒症，发病率（$2p$）为 1.063×10^{-4}。已知本病的选择系数 $s=0.80$，故基因突变率为 $\nu=sp=0.80\times1/2\times1.063\times10^{-4}=4.25\times10^{-5}$。

（三）选择对隐性基因的作用

1. 致病基因频率的改变　群体中基因型 AA 和 Aa 个体的频率分别为 p^2 和 $2pq$，其表型正常，适合度为1；而基因型 aa 个体的频率是 q^2，表型不正常，适合度为 $1-s$。经一代选择后，基因 a 频率的改变为：

$$\Delta q=\frac{-sq^2(1-q)}{1-sq^2}\approx -sq^2(1-q) \tag{7-18}$$

式7-18表明，当 q 值较大（此时 s 值通常较小）时，q 的改变较大。但在人类群体中，q 值往往很低，每代基因 a 频率的降低大致为 $sq^2(1-q)$，因此隐性致病基因的频率降低速率往往很缓慢（见表7-9）。

表7-9 显性完全，选择对隐性纯合子不利时，基因频率 q 的改变

	AA	Aa	aa	合计
原来频率	p^2	$2pq$	q^2	1
适合度	1	1	$1-s$	
选择后	p^2	$2pq$	$q^2(1-s)$	$1-sq^2$
相对频率	$\frac{p^2}{1-sq^2}$	$\frac{2pq}{1-sq^2}$	$\frac{q^2(1-s)}{1-sq^2}$	1
基因 a 频率原来为 q	选择后为：$\frac{q(1-s)}{1-sq^2}=\frac{pq+q^2(1-s)}{1-sq^2}$			
基因 a 频率的改变：	$\Delta q=\frac{q(1-s)}{1-sq^2}-q=\frac{-sq^2(1-q)}{1-sq^2}\approx -sq^2(1-q)$			

2. 基因突变率的确定　尽管隐性致病基因的频率在群体中降低得很慢，但其最终趋势仍是从群体中消失。同时，新的隐性突变会不断产生，弥补因选择作用而淘汰的隐性致病基因。因此，一个隐性遗传病发病率相对稳定的群体，某种程度上可以视为突变和选择作用实现了平衡。因此有：

$$\mu=\Delta q=sq^2(1-q)\approx sq^2 \tag{7-19}$$

据此，知道 s 和 q 值后，就可推算突变率。例如苯丙酮尿症是一种隐性遗传病，在我国人群中的发病率(q^2)为 6.0×10^{-5}。已知该病患者的选择系数为 0.85。据此求得基因突变率为 $\mu=5.1\times10^{-5}$。如果 $s=1$，即隐性纯合体致死时，突变率就正好等于发病率。

（四）选择对 X 连锁基因的作用

X 连锁基因所决定的性状在男性中显示出来，但有 2/3 的 X 连锁隐性基因存在于杂合子女性中，这些女性在表型上正常，不受选择的作用。因此，选择对 X 连锁隐性基因的作用强于对常染色体隐性基因的选择，但是不及对常染色体显性基因的选择。

A 型血友病是由 X 连锁隐性基因决定的性状，几乎所有的患者都是男性。已知群体中血友病基因的频率是 q，则 q 也是男性血友病的发病率。由于选择的作用，每一代男人中有频率为 sq 的血友病基因被淘汰，在平衡群体中则必须由相同数量的正常基因的突变来平衡，同时，由于男性中每有一个血友病基因被表达，预示着还有 2 个血友病基因存在于杂合子女性中，因此：

$$\mu=\frac{1}{3}sq \tag{7-20}$$

据调查，A 型血友病的男性发病率(q)为 8.0×10^{-5}，$s=0.75$，据此求得其基因突变率为 $\mu=2.0\times10^{-5}$。

选择分别对显性、隐性、X-连锁致死基因的选择效率的结果见图 7-2。图中所示的选择是“完全”的。其起始频率均为 0.25(引自文献 8)。

需要指出，选择是一个复杂的生物过程，一个基因型对另一个基因型的选择优势，既取决于基因和环境因素间的相互作用，也取决于不同座位基因间的相互作用。正如 S Wright 曾强调的，自然选择不是作用在基因之间，而是作用在个体之间。

三、群体中的平衡多态现象

群体中的多态现象(polymorphism)是指群体中同一基因座上有两个或两个以上的等位基因同时存在，并且其中频率最低的等位基因频率也远远高于仅靠突变所能维持的基因频率。譬如人群中的 ABO 血型，MN 血型，等等。

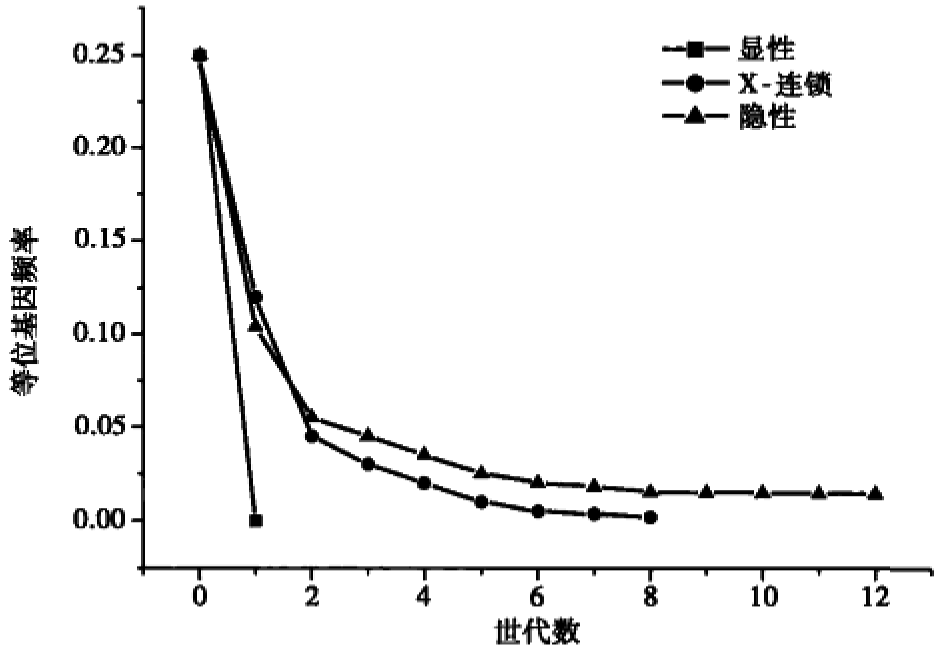


图 7-2 选择分别对显性、隐性、X-连锁致死基因的选择效率的结果

产生多态现象的原因较多,有些至今也未有定论。一般说来,大致有以下几种原因可以造成群体的多态现象。

(一) 平衡选择现象

突变是多态现象的最初起源,但极低的突变频率本身无法维持群体的多态。前面介绍的几种常见的选择作用也无法维持群体的多态。然而,当杂合子的适合度比两个纯合子都高,即所谓超显性(overdominance)现象存在时,则选择作用就可保持群体的多态现象。

设杂合子 *Aa* 的适合度为 1,纯合子 *AA* 和 *aa* 的适合度分别为 $1-s_1$ 和 $1-s_2$,其中 $s_1>0, s_2>0$。那么经一代选择后,基因 *a* 频率的改变见表 7-10。

表 7-10 杂合子的适合度比两个纯合子都高时,经一代选择后基因频率的改变

	AA	*Aa*	*aa*	合计
原来频率	p^2	$2pq$	q^2	1
适合度	$1-s_1$	1	$1-s_2$	
选择后	$p^2(1-s_1)$	$2pq$	$q^2(1-s_2)$	$1-p^2s_1-q^2s_2$
相对频率	$\frac{p^2(1-s_1)}{1-p^2s_1-q^2s_2}$	$\frac{2pq}{1-p^2s_1-q^2s_2}$	$\frac{q^2(1-s_2)}{1-p^2s_1-q^2s_2}$	1
基因 *a* 频率原来为 *q*		选择后为:$\frac{q(1-qs_2)}{1-p^2s_1-q^2s_2}$		
基因 a 频率的改变:		$\Delta q=\frac{q(1-qs_2)}{1-p^2s_1-q^2s_2}-q=\frac{pq(ps_1-qs_2)}{1-p^2s_1-q^2s_2}$		

如果选择的作用导致群体维持基因 *a* 的频率保持不变,则有

$$\Delta q=\frac{pq(ps_1-qs_2)}{1-p^2s_1-q^2s_2}=0 \tag{7-21}$$

解得 $\Delta q=0$ 的条件为:①$p=0$;②$q=0$;③$q=s_1/(s_1+s_2)$。不考虑前两种情况($p=0$ 或 $q=0$ 条件下,该基因无群体多态现象),第③种情况表明:达到平衡多态的基因 *a* 的频率由两种纯合子的选择系数所决定。这种现象也称为平衡选择多态。

最著名的例子是镰形红细胞性状。有镰形红细胞性状的杂合子个体($\beta^A\beta^S$)对恶性疟原虫的抵

抗能力较强。因此在赤道非洲的某些群体中,镰形红细胞贫血患者的频率可高达4%,群体中β^S的基因频率高达0.2,而该病的患者几乎都在成年前死去,基本没有后代,即$1-s_2=0$,由$q=s_1/(s_1+s_2)$可求得正常纯合体在该环境条件下的适合度为$1-s_1=0.75$。这个例子也说明:①不存在在任何环境中都表现为最高适合度的基因型,环境的改变必然导致基因型适合度的改变;②只要杂合子的适合度比正常纯合子稍有增加,就可补偿因隐性纯合子的致死而丧失的隐性基因,使群体维持多态。

(二) 中性突变-随机漂变学说与多态现象

尽管杂合子优势可以解释一部分遗传多态现象,但绝大多数的遗传多态现象,譬如ABO血型,MN血型等,无法由经典群体遗传学中获得圆满解释。在DNA水平上,绝大多数新突变与群体内先存的等位基因不同,因此由一对中性等位基因之间的正向突变与逆向突变来实现基因频率平衡的经典假说已被证明是不切实际的。

在对分子水平上更为丰富的多态现象进行大量综合分析的基础上,Kimura等提出了中性突变-随机漂变学说,解释分子水平上的多态现象,该学说的要点是:

1. 在分子水平上,许多突变是有害的,相当数量的突变是中性或近中性的;仅有很小一部分突变是有利的。

2. 自然选择仅仅是一种保存有利突变和消灭有害突变的进化过程。

3. 大部分新突变都将消失,少量新突变的固定依赖于随机漂变(random drift)。

4. 在分子水平上,群体存在巨大的遗传变异,造成丰富的遗传多态。

5. 群体的多态性最先仅仅由突变产生。最初的突变体对后代有着最显著的影响,称为"奠基者效应"(founder effect)。

该理论中有关基因的随机漂变和固定的基本概念,将在第四节作进一步介绍。

(三) 群体中维持遗传多态的其他机制

群体内遗传多态的维持机制是个非常复杂的问题。除了上述杂合子优势的"平衡选择现象"和"遗传漂变"外,在自然选择的作用下,还有许多因素可以维持一个群体的遗传多态。以下列举几个较常见的维持群体中遗传多态的机制:

1. 环境的空间分布差异。选择与适应相互作用的结果会导致某种等位基因适宜在某种环境条件下生活,而其他形式的等位基因适宜在另一类环境下生活;

2. 上位性效应,包括非等位基因间的互作、修饰、协同进化,等等。有些等位基因很适应某种特定的遗传背景,而不能适应其他的遗传背景;

3. 频率依赖选择。有的等位基因当其频率很低时,常可表现出某种选择优势,从而促使该等位基因频率得到某种程度的恢复。

第四节 影响群体基因频率的因素(二):遗传漂变和迁移

一、遗传漂变

前面讨论Hardy-Weinberg平衡律以及突变和选择对群体基因频率的影响时,研究对象通常是一个很大的群体,理论上称之为无限群体,并且群体内进行完全的随机婚配。然而,实际的人类自然群体均为有限群体,每对父母生育的子女数也极为有限,因此,即使缺乏自然选择和人为选择的作用,群体下一世代也不可能完全重现上一世代的基因和基因型频率。由于配子间的结合可能为随机抽样的结果,必然会导致随机偏离,使下一代的基因频率在上一代基因频率的附近随机摆动。摆动的幅度一般取决于群体的有效容量和群体的基因频率。这种摆动将一代代持续下去,并且由于下一世代的基因频率只依赖于上一世代的群体遗传参数,而与更久远的世代的群体状态并没有直接关系,因此,小群体中这种世代间的基因频率的随机变化最终会导致一个等位基因的固定或丢失,形成某个等位基

因的纯合子亚群体。这种由于群体样本容量的有限性和基因在世代传递中的随机抽样作用造成的基因频率的随机波动被称为随机遗传漂变(random genetic drift)。任何一个隔离的人类自然群体,如果与其他群体间没有基因的交流,就会经历这样的过程。

遗传漂变的方向无法肯定,但漂变的范围却可以估计。假设有一群体,其有效的群体容量为N(对通常的人类自然隔离群体而言,N值一般可视为实际群体中个体数量的1/3),对于常染色体的某一基因座,共有2N个基因。设该座位的两个等位基因分别为A和a,基因频率分别为p和q($q=1-p$),则下一世代基因频率改变的方差为

$$Var_{\Delta p}=\frac{pq}{2N} \tag{7-22}$$

由上式可计算出各种不同容量的群体中的群体基因频率的标准差,见表7-11。

假定有效群体大小(N)为50,且世代之间维持不变。在初始基因频率$p=q=0.50$,不考虑突变和选择的作用以及随机婚配的假定下,子代群体的基因频率有68%的可能出现在0.45~0.55之间,也有5%的可能出现在0.40~0.60之外。

表7-11 遗传漂变导致的群体基因频率的抽样误差($p=q=0.50$)

	群体大小			
	10	50	100	200
$\sqrt{\frac{pq}{2N}}$	0.11	0.05	0.04	0.03

因此,小群体里基因频率的随机漂变可能使选择中性或接近中性的基因在较少的世代数内被固定或消失,有时甚至也使选择上不太有利的基因频率增加乃至固定,或者使选择上较有利的基因频率减少乃至消失。

遗传漂变的另一种形式,发生在从一个较大群体中分离出一个小群体并在此基础上逐渐发展起来的群体中。这种剧烈的"漂变"称为"奠基者效应"(founder effect)。美洲印第安人缺乏B型血型,很有可能是由于约20000年前,从亚洲迁去的这些印第安人的祖先群体中就不含B型等位基因,或在最初的几个世代因随机漂变而导致该等位基因从群体中消失。

遗传漂变的另一个著名例子是,东卡罗林群岛的Pingelap人中,一种常染色体隐性遗传的先天性失明症有相当高的发病率。在1780~1790年间,一次飓风袭击了Pingelap岛,造成岛上居民的大量死亡,岛上只留下9个男人和数目不多的女人。推测起来,最初留下的少数人中可能含有较高频率的"先天性失明"基因,因此尽管该基因在选择上处于不利位置,奠基者效应却导致在以后的数个世代内持续了这种高频率。最终导致到1970年,该岛的1500余人中,盲人的比例竟高达4%~10%之多。

在可视为无限大的配子库中,随机抽取有限数量样本的结果是:群体中大多数新的突变,即使不受选择压力的作用,也将很快从群体中消失。假设有一个体是一个新突变的杂合子,则突变的基因将有两种命运:①由于某种原因,这一个体未能留下后代,则该突变将从群体中消失;②该个体留下了后代,但该突变等位基因仍有可能无法传递给子女,以此类推,每一代都会遇到该等位基因是否会丢失的问题。事实上,对于一个有效容量为N的群体,一个不受选择压力作用的新的突变从群体中丢失的概率为$1-\frac{1}{2N}$。另一方面,新的突变如果能一直保存在群体中不被丢失,最终将遍及整个群体直至被固定。因此,群体愈小,新的突变等位基因被固定的概率就愈大;反之,群体愈大,新的突变等位基因从群体中消失的概率就愈大。即使是一个选择上稍微有利的新的突变,在产生后的最初几个世代里也极有可能因遗传漂变而消失。当然,"命运"会稍许好些。譬如,如果突变杂合子个体表现出比正常纯合体高出1%的适合度,则该突变等位基因从群体中消失的概率,是纯粹因随机漂变导致在群体中消失的概率的98%。

二、迁移和混合

群体遗传变异的另一个重要来源是群体间的基因交流,即群体的混合行为。混合行为通常导致新群体的基因频率介于两个初始群体之间。混合方式大致可划分为迅速混合和渐进混合两类,不同

的混合方式对新的群体基因频率的影响也有所不同。

(一) 迅速混合方式

迅速混合方式，即两个种族、生活习惯和文化背景等差异较大的群体，由于战乱或人群的迁徙等原因混居在一起，并在较短的世代内融合成一个新的群体。设混合群体(admixed population)Z由比例为m的X群体和比例为$1-m$的Y群体在较短时间内混合而成。p_x、p_y分别为某一基因座位上某等位基因在初始群体X、Y的频率，则新的混合群体中，该等位基因的频率p_z为：

$$p_z = mp_x + (1-m)p_y \tag{7-23}$$

故混合群体的基因频率取决于初始群体的基因频率和群体的混合比例。特别地，当$m=0.5$时，混合群体的基因频率为初始群体基因频率的均值。

(二) 渐进混合方式

假定某个群体在每一世代都有一定比例个体融入另一群体，则接受外来基因的群体的基因频率将逐渐发生变化。设p_t为接受群体(recipient population)在t世代某等位基因的频率，m为供应群体(donor population)每代输入接受群体的基因占接受群体基因的比例(贡献率)，p'为供应群体该等位基因的频率。则下一世代接受群体的该等位基因频率p_{t+1}为：

$$p_{t+1} = (1-m)p_t + mp'$$

由此引起的频率变化为：

$$\Delta p = p_{t+1} - p_t = m(p' - p_t) \tag{7-24}$$

因此，接受群体世代间的频率变化取决于贡献率m和两个群体等位基因频率之差。如果供应群体和接受群体的基因频率差别较大，且贡献率m也较大的话，则接受群体的基因频率变化率可以很大。美洲新大陆被发现以来，最初作为奴隶被贩运到美洲的黑人群体就一直接收来自白人群体的基因。因此，一种等位基因如果只在欧洲人群发现，却不存在于非洲人群中，就可以通过测定目前美洲黑人群体中该等位基因的频率来确定m值。

如果混合速率不是很大，则自混合发生以来的每代混合率m_i的和M(即总的供应群体基因贡献率)与若干代后接受群体的基因频率的总变化有关。即

$$p_t - p_0 = \Delta p_{total} = M(p' - p_0)$$

所以

$$M = \frac{\Delta p_{total}}{p' - p_0} \tag{7-25}$$

例如Duffy blood血型系统中的F_y^a等位基因在非洲人群中不存在，但在美国乔治亚洲的白人人群中却有0.42的高频率。同时，该州黑人人群中该等位基因的频率为0.046。因此，白人人群对该州黑人人群总的基因贡献率为

$$M = \frac{\Delta p_{total}}{p' - p_0} = \frac{0.046-0}{0.42-0} = 0.1095$$

三、遗传漂变、选择及基因交流作用的相对有效性

影响群体遗传结构的因素很多，突变和选择的效应依赖于群体容量的大小；遗传漂变的作用与群体大小N成反比；而确定性模型反映的改变群体遗传结构的力量又依赖于迁移率m，突变率μ和选

择系数 s；等等。

大体上，我们可以这样认为，如果 $m \geq 1/N$ 或 $\mu > 1/N$，则群体混合或突变的作用占主导地位；如果 $s \geq 1/N$，则选择的作用不可忽略。如果选择系数过小，或群体容量过小，则该突变在群体中可视为"选择中性"。在一个很小的群体内，漂变的作用通常大于选择的作用。

就人类群体而言，几乎人类历史上的所有时间，人类群体的容量都很小，只是在最近的几百代里有了明显的快速增长。因此，我们可以认为，他们之所以能在一个较小的亚群或封闭群体内达到较高的基因频率，主要依赖于遗传漂变的作用。

第五节 近婚系数

Hardy-Weinberg 平衡律是基于"随机婚配"假定条件下导出的，然而我们必须仔细区别"随机婚配"的两方面含义。

对那些不左右配偶选择的性状(如与外观、习性、气味、智力状况或疾病等无关的可遗传性状)而言，随机婚配的假设是可以成立的。譬如人们通常不大会在意配偶的血型。而将血型作为择偶的一个标准，则更为罕见。对这类基因而言，婚配是"随机"的。

然而，对许多可遗传的性状而言，婚配不是随机的或不完全是随机的，当人类出现分群时，群体内的婚配机会通常总是高于群体间的婚配。如果某一基因座上群体间等位基因频率差别较大，即使群体内满足 Hardy-Weinberg 平衡，整个人类群体也无法实现这种平衡(见表 7-12)。

表 7-12 MN 血型的观测频率和期望频率的比较[5]

	观测频率			期望频率		
	M/M	M/N	N/N	M/M	M/N	N/N
爱斯基摩人	0.836	0.156	0.009	0.834	0.159	0.008
埃及人	0.278	0.489	0.233	0.274	0.499	0.228
中国人	0.332	0.486	0.182	0.331	0.488	0.181
澳洲土著人	0.024	0.304	0.672	0.031	0.290	0.679

非随机婚配比较常见的类型有：婚配倾向于发生在有一定亲缘关系个体间(近交，inbreeding)，或倾向于远离亲缘关系的个体间(enforced outbreeding 或 negative inbreeding)。

此外，个体间不是根据亲缘关系的远近，而是根据肤色、身高、生活习性、智力状况等因素的相似程度来择偶，称为选型婚配(assortative mating)。选型婚配的结果将导致群体内纯合度的增加。

一、常染色体近婚系数

(一) 近婚系数的定义

与选型婚配一样，近亲婚配(consanguineous marriage)的结果也会导致群体纯合度的增加。两个个体间存在亲缘关系，则他们至少享有一个共同祖先。他们的某个等位基因有可能来源于同一祖先相同的 DNA 序列。如果这两个个体进行婚配，他们的后代有可能出现这样的情形，即同一座位的一对等位基因是同一祖先同一序列的两个拷贝(identical by descent，IBD，后裔同源)。两个婚配个体间的亲缘关系愈近，后代基因纯合(homozygosity by descent)发生的概率就愈高。这种后代发生同一祖先同一基因纯合的概率，就定义为近婚系数(inbreeding coefficient)，通常用符号 F 表示。

（二）表兄妹婚配的常染色体近婚系数

设 P_1 的基因型为 A_1A_2，P_2 的基因型为 A_3A_4。从图 7-3 中可知，P_1 把基因 A_1 传给 B_1 的概率是 $\frac{1}{2}$，B_1 得到基因 A_1 将其传给 C_1 的概率是 $\frac{1}{2}$，而 C_1 得到基因 A_1 后再将其传给 S 的概率还是 $\frac{1}{2}$。这样 S 从 P_1 经 B_1、C_1 获得基因 A_1 的概率即为 $\left(\frac{1}{2}\right)^3$；同理，$S$ 从 P_1 经 B_2、C_2 获得基因 A_1 的概率也为 $\left(\frac{1}{2}\right)^3$；故 S 的基因型为 A_1A_1 的概率为 $\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^3=\left(\frac{1}{2}\right)^6$。同样，$S$ 的基因型为 A_2A_2、A_3A_3 或 A_4A_4 的概率也为 $\left(\frac{1}{2}\right)^6$。因此，$S$ 的近婚系数为 $F=4\times\left(\frac{1}{2}\right)^6=\frac{1}{16}$。

（三）二级表兄妹婚配的常染色体近婚系数

如图 7-4 所示，二级表兄妹（从表兄妹）婚配的情况下，基因的传递过程比表兄妹婚配又增加了 1 步，故其近婚系数为 $F=4\times\left(\frac{1}{2}\right)^8=\frac{1}{64}$。

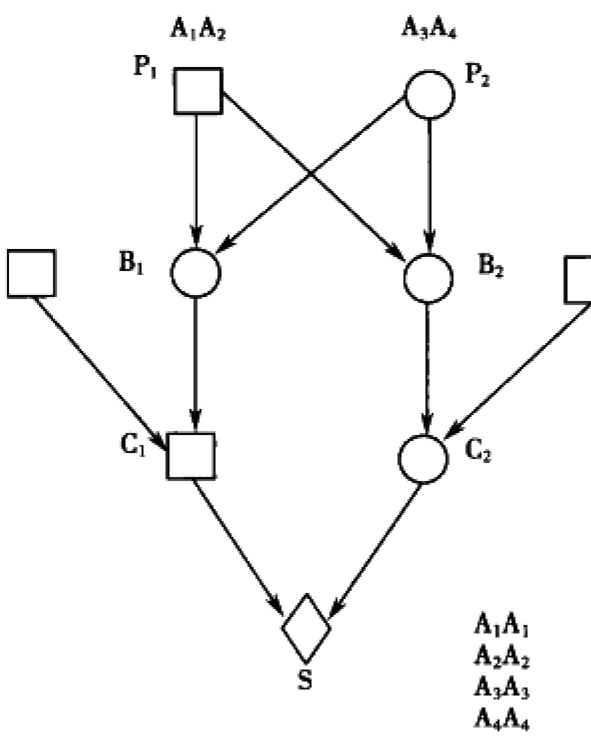


图 7-3　表兄妹婚配中基因传递图解

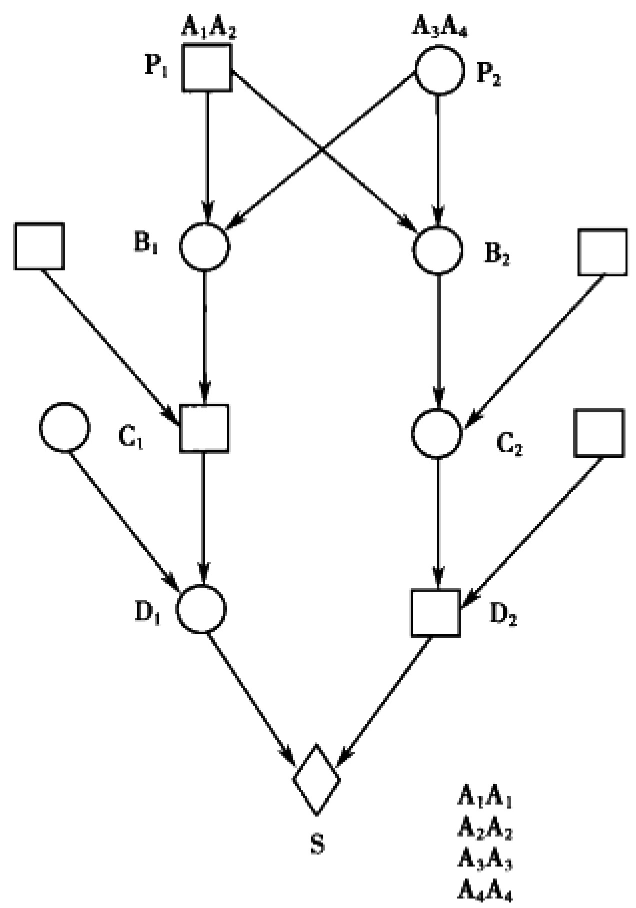


图 7-4　二级表兄妹婚配中基因传递图解

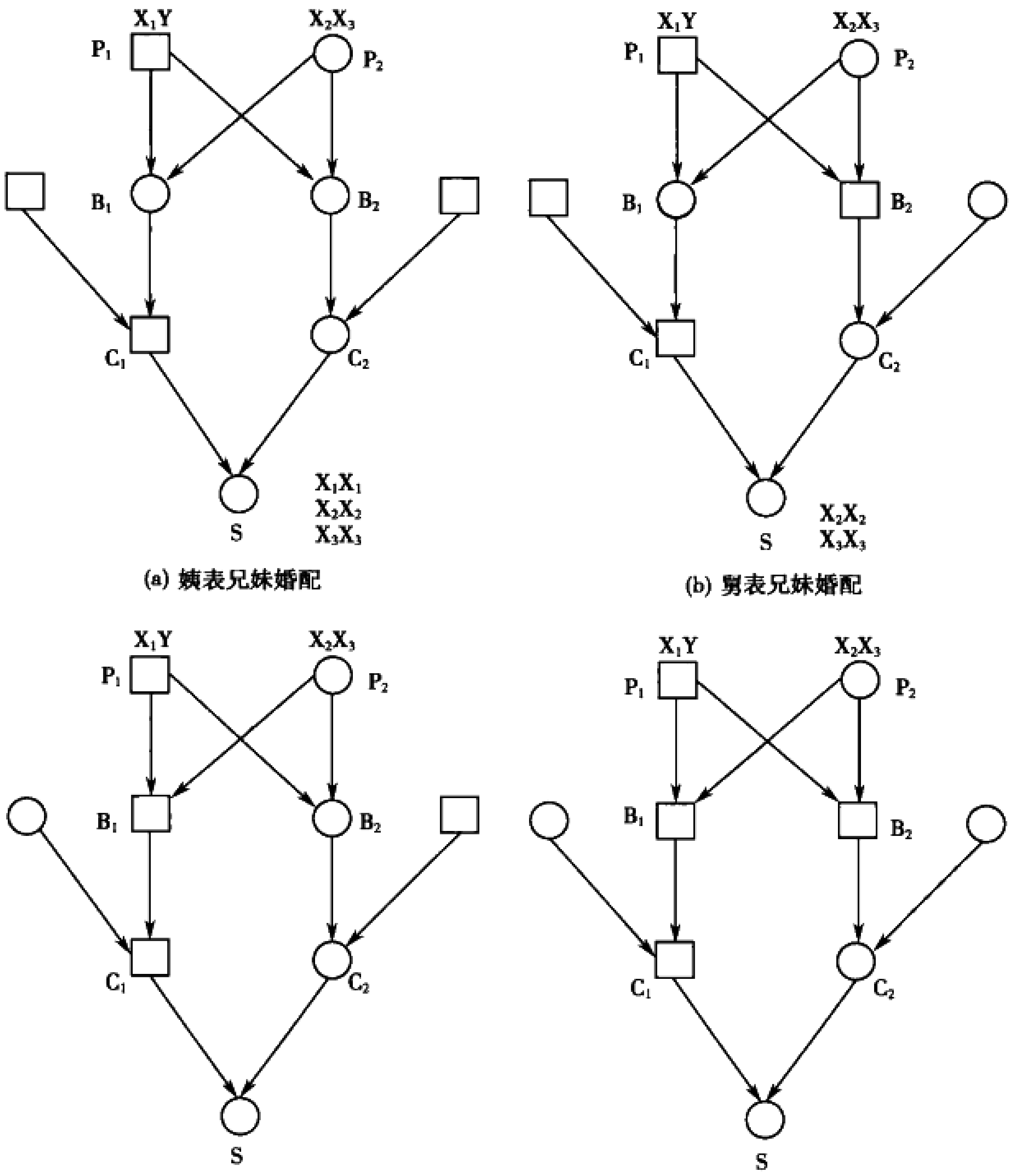
（四）其他形式近亲婚配的近婚系数

通过类似的基因传递过程分析，可以得到其他形式近亲婚配的近婚系数，这里不再一一分析。其主要结果为：①舅甥女（或姑侄）间的近婚系数为 $\frac{1}{8}$；②表舅甥女（或堂姑侄）间的近婚系数为 $\frac{1}{32}$；③半表兄妹（只有一个共同的祖先）的近婚系数为 $\frac{1}{32}$；④半从表兄妹间的近婚系数为 $\frac{1}{128}$。

二、X 连锁基因的近婚系数

对常染色体上的基因而言，父母为近亲结婚时，对儿子和女儿的影响程度相同。但对于 X 染色体上的基因，情况则有所不同。由于男性只有一条 X 染色体，基因不存在纯合性问题，因此当父母是近亲结婚时，对儿子无影响。同时，从传递特点来看，男性的 X 连锁基因一定传给他的女儿，所以传

递概率为 1；相反，男性的 X 连锁基因不可能传递给他的儿子，所以传递概率为 0。因此，在计算有关 X 连锁基因的近婚系数时，只计算女儿的 F 值。图 7-5（a）～（d）分别是姨表兄妹、舅表兄妹、姑表兄妹和堂表兄妹婚配的 X 连锁基因的传递图解。



(a) 姨表兄妹婚配　(b) 舅表兄妹婚配　(c) 姑表兄妹婚配　(d) 堂兄妹婚配

图 7-5　表兄妹婚配的 X 连锁基因传递图解

姨表兄妹婚配中，基因 X_1 由 P_1 经 B_1、C_1 传至 S，只计 1 步（B_1 传至 C_1）；基因 X_1 经 B_2、C_2 传至 S 需 2 步（B_2 传至 C_2 再传至 S）。故 S 为 X_1X_1 的概率为$\left(\frac{1}{2}\right)^3$。基因 X_2 从 P_2 经 B_1、C_1 传至 S，需计两步，基因 X_2 从 P_2 经 B_2、C_2 传至 S，却需计 3 步，所以 S 为 X_2X_2 的概率为$\left(\frac{1}{2}\right)^5$；同理 S 为 X_3X_3 的概率也是$\left(\frac{1}{2}\right)^5$。因此，姨表兄妹婚配的近交系数为$\left(\frac{1}{2}\right)^3+2\times\left(\frac{1}{2}\right)^5=\frac{3}{16}$。

舅表兄妹婚配，基因 X_1 从 P_1 传至 B_2 时中断，不能形成 X_1X_1。基因 X_2 从 P_2 经 B_1、C_1 向 S 传递，只需计为传递 2 步；基因 X_2 从 P_2 经 B_2、C_2 向 S 传递，也只需计为传递 2 步；所以 S 为 X_2X_2 的概率为$\left(\frac{1}{2}\right)^4$；同理 S 为 X_3X_3 的概率也是$\left(\frac{1}{2}\right)^4$，故舅表兄妹婚配的近交系数为$2\times\left(\frac{1}{2}\right)^4=\frac{1}{8}$。

姑表兄妹婚配，基因 X_1 从 P_1 传至 B_1 时中断，基因 X_2 和 X_3 从 P_2 经 B_1 传至 C_1 时中断，故近交系数为0。

堂表兄妹婚配，基因 X_1 从 P_1 传至 B_1 时中断，基因 X_2、X_3 从 P_2 经 B_1 传至 C_1 时中断，故近交系数也为0。

因此，仅就X连锁基因来看，姨表兄妹婚配的近婚系数大于舅表兄妹婚配；姑表兄妹和堂表兄妹婚配的近婚系数均为0。

三、近亲婚配的遗传学效应

已知有亲缘关系个体间的婚配系数和群体中某个隐性致病基因的基因频率 q，就可推断近亲婚配生育隐性纯合子的概率，并据此估计近亲婚配的有害程度。

（一）近亲婚配的隐性纯合概率

近亲婚配的危害，主要表现在隐性遗传病纯合体患者的频率增加。有两种原因可以导致近亲婚配的子女是隐性致病基因纯合体（aa）：①从共同祖先传递而来，形成纯合体 aa 的概率为近婚系数 F 与隐性致病基因的频率 q 的乘积 Fq；②由不同祖先分别传来，形成纯合体的概率为 $(1-F)q^2$。①、②两种情况合计，近亲婚配产生隐性纯合体的概率为：

$$Fq+(1-F)q^2 = Fq+q^2-Fq^2 = q^2+Fq(1-q) = q^2+Fpq \tag{7-26}$$

因此，对常染色体上的基因而言，近亲婚配的结果是导致其产生隐性纯合体的概率在随机婚配产生隐性纯合体的概率 q^2 的基础上再增加 Fpq，其中 p 为显性基因的基因频率。故表兄妹婚配导致隐性纯合的概率为 $q^2+\frac{pq}{16}$；从表兄妹婚配导致隐性纯合的概率为 $q^2+\frac{pq}{64}$。

（二）近亲婚配导致隐性纯合的相对风险（β）

式7-26表明，群体中隐性致病基因的基因频率愈低，近亲婚配导致该病发病的频率也愈小。但从另一个角度说，自然群体由于近亲婚配导致隐性纯合的概率，与随机婚配导致隐性纯合的概率的相对比例却增大，为了说明这个问题，我们定义 β 为近亲婚配导致隐性纯合的相对风险。

$$\beta = \frac{q^2+Fpq}{q^2} = 1+\frac{Fp}{q} \tag{7-27}$$

式7-27表明，近亲婚配导致隐性纯合的相对风险总是大于或等于随机婚配，近婚系数愈大，群体中致病基因频率愈低，则近亲婚配导致隐性纯合的相对风险愈高。所以，愈是罕见的隐性遗传病，病儿出自近亲婚配的概率愈大。表7-13给出不同近婚系数和群体基因频率条件下随机婚配和近亲婚配导致隐性纯合的概率及其隐性纯合的相对风险。

表7-13 随机婚配和近亲婚配常染色体隐性纯合的频率及近亲婚配的相对风险

基因频率（q）	随机婚配（q^2）	表亲婚配 $\left(q^2+\frac{pq}{16}\right)$	表亲婚配的相对风险 β_1	从表亲婚配 $\left(q^2+\frac{pq}{64}\right)$	从表亲婚配的相对风险 β_2
1.0×1.0^{-1}	1.0×1.0^{-2}	1.5625×1.0^{-2}	1.5625	1.1406×1.0^{-2}	1.1406
5.0×1.0^{-2}	2.5×1.0^{-3}	5.4688×1.0^{-3}	2.1875	3.2422×1.0^{-3}	1.2969
1.0×1.0^{-2}	1.0×1.0^{-4}	7.1875×1.0^{-4}	7.1875	2.5469×1.0^{-4}	2.5469
1.0×1.0^{-3}	1.0×1.0^{-6}	6.3438×1.0^{-5}	63.4375	1.6609×1.0^{-5}	16.6094
1.0×1.0^{-4}	1.0×1.0^{-8}	6.2594×1.0^{-6}	625.9375	1.5723×1.0^{-6}	157.2343

需要指出的是，表 7-13 只是比较近亲婚配与随机婚配产生隐性纯合体的相对风险，并不代表自然群体中隐性纯合体出自近亲婚配的实际比例。此外，随着人类社会文明的发展，人类活动和择偶范围不断扩大，医学遗传学知识也越来越普及，加上婚姻法的限制，近亲婚配的机会将不断减少，自然群体中隐性纯合体出自近亲婚配的比例也将持续下降。

（三）近亲婚配与群体遗传负荷

人类群体中因纯合而有害的等位基因并不少见，这种由于基因纯合导致群体适合度降低的现象，称为遗传负荷（genetic load）。遗传负荷一般用群体中每个个体平均所携带的致死基因或有害基因的数量来衡量。一般的估计倾向于认为，一个人可能带有 4～8 个有害基因。近亲婚配的一个明显效应，就是使纯合子的频率增加。因此近亲结婚会增加群体的遗传负荷，从而导致群体适合度的降低。

四、瓦赫伦（Wahlund）效应

前面的讨论将近婚效应作为自然群体非随机婚配的主要原因。事实上，人类自然群体还存在另一类“非随机”婚配方式：婚姻通常更倾向于选择同一种族、具有相同或相近的文化背景、宗教信仰、生活习俗的配偶，或者说，人类的婚姻半径并不十分大，在较为封闭的区域这种现象更为明显。如果一个大的群体事实上划分为若干个小的群体（这在自然人群中是较普遍的现象），则尽管在各个小群体内都是随机婚配，但在总的群体上，依然可以观察到群体基因型频率偏离 Hardy-Weinberg 平衡定律的现象（即总体上可视为一定程度的近婚群体）。这种现象称为瓦赫伦效应（Wahlund effect）。

瓦赫伦（Wahlund）效应的理论分析：假定一个大群体可划分为 k 个亚群体，等位基因 A 在第 i 个亚群体内的基因频率为 p_i，等位基因 a 在第 i 个亚群体内的基因频率为 q_i；假定所有的亚群体的有效群体容量相等，且所有亚群体内的个体均为随机婚配，则大群体等位基因 A 的平均频率为$\bar{p}=\sum p_i/k$，等位基因 a 的平均频率为$\bar{q}=1-\bar{p}$。如果整个大群体内实现随机婚配，则群体 AA 基因型频率为$(\bar{p})^2$；然而，当大群体划分为 k 个有效容量相等的亚群体且亚群体内实现随机婚配时，我们所观测到的 AA 基因型平均频率为：

$$\begin{aligned}\sum p_i^2/k &= \sum(p_i-\bar{p}+\bar{p})^2/k=\sum((p_i-\bar{p})^2-2\bar{p}(p_i-\bar{p})+(\bar{p})^2)/k\\ &=\sum(p_i-\bar{p})^2/k+(\bar{p})^2=\mathrm{Var}(p)+(\bar{p})^2\end{aligned} \tag{7-28}$$

式中 Var(p)为大群体内等位基因 A_1 的方差。同理可得：

$$\sum 2p_iq_i/k = 2\bar{p}\,\bar{q}-2\mathrm{Var}(p) \tag{7-29}$$

$$\sum q_i^2/k = \bar{q}+\mathrm{Var}(p) \tag{7-30}$$

依据方差的性质（$\mathrm{Var}(p)\geqslant 0$），只要大群体内所有的亚群体基因频率不是全部相等，就可得到如下结论：

1. 大群体分为若干个亚群体后，整个群体的纯合子将增加，杂合子将减少。其效应类似于一定程度的近亲婚配；
2. 纯合度增加及杂合度减少的程度，取决于亚群体间基因频率的方差的大小；
3. 对于复等位基因，上述原则仍成立。

第六节 优生运动的误区

一、“Eugenics”一词不再适于在科学文献中使用

Eugenics 一词源于希腊文，中文译为“优生学”，原意是 well-born 或 healthy birth，即“生出一个健康的孩子”，这本来是人类的美好愿望。但随着遗传学的蓬勃发展，某些人类遗传学家开始设想利用

遗传学手段来改良人种，20 世纪 30～40 年代在西方掀起了盛极一时的优生运动。尤其在德国，一场旨在消灭"劣生者"的优生运动的泛滥，最终竟导致了惨绝人寰的大屠杀！历史已将 Eugenics 一词与希特勒对所谓"劣生者"强制绝育并导致种族灭绝的行为联系在一起，在概念上被指称为"由国家强加于个人生育的社会规则"，因此，该词不再适于出现在科学文献中。

纳粹德国遗传学和优生学留给后人的教训是深刻的。一个人，不管他属于什么种族，也不管他身体是否健康、智力是否正常，在道德和法律上与其他人都是平等的。任何组织或个人都无权剥夺他生命、健康、结婚和生育的权力。因此，政府或法律对婚育的限制应该是最低限度的，例如对近亲通婚的限制，或者在人口爆炸时对生育数量的暂时限制。

基于此，1998 年在北京召开的第 18 届国际遗传学大会上，有关"遗传学的伦理、法律和社会含义"及"优生学的科学和伦理"的讨论非常热烈。最后达成了八点共识：

1. 建立在追求"有利"和"无害"基础上的许多伦理原则为大多数国家所共同遵守。这些原则的应用可有许多不同的方式。

2. 新的遗传学技术应该用于为个人提供可靠的信息，在此基础上作出个人生育选择，而不应作为制订强制性公共政策的工具。

3. 知情选择应该是有关生育决定的一切遗传咨询和建议的基础。

4. 遗传咨询应该有利于夫妇和他们的家庭，人为限制生育的措施对有害性等位基因在人群中的发生率影响极小。

5. "Eugenics"这个术语以各种不同的方式被滥用，使其已不再适于在科学文献中使用。

6. 在制定关于健康的遗传方面的政策时，应该在各个层次进行国际和学科间的交流。

7. 关注人类健康的遗传方面的决策者有责任征求正确的科学意见。

8. 遗传学家有责任对医师、决策者和一般公众进行遗传学及其对健康的重要性的教育。

二、限制遗传疾病患者生育不能清除遗传疾病

有人天真地认为，只要不让有遗传疾病的人生育，就可以减少遗传疾病患者，最终从群体中剔除该遗传疾病。然而分析表明，当隐性致病突变基因的频率很低时，人为的选择作用对群体突变基因频率的改变影响极小（详见本章第三节式 7-18）。各种选择系数（s）条件下有害隐性基因频率变化 Δq 所需要的世代数见表 7-14。

表 7-14　在不同选择系数作用下，有害隐性基因频率改变所需的世代数

基因频率 q 的改变从 q_0 到 q_n		$s=1$（致死）	不同 s 值所需的世代数		
			$s=0.5$	$s=0.1$	$s=0.01$
0.99	0.50	1	11	56	559
0.50	0.10	8	20	102	1020
0.10	0.01	90	185	924	9240
0.01	0.001	900	1805	9023	90 231
0.001	0.0001	9000	18 005	90 023	900 230

表 7-14 表明，基因频率从 0.99 淘汰到 0.10 是比较快的，对于几乎所有选择系数都是这样。可是基因频率再进一步降低就相当缓慢，例如，基因频率从 0.001 降低到 0.0001，即使纯合致死，也需要 9000 代。

对频率较低的隐性个体的选择无效的原因在于，绝大多数隐性有害基因存在于杂合体中，自然选择对它们不起作用。隐性患者与携带者的比例为 $q^2:2pq=q:2p$，所以隐性基因在群体中出现的频率

愈低,它存在于杂合体中的相对机会就愈高。例如 $q = 0.01$,隐性纯合体频率(q^2)是0.0001,而杂合体频率($2pq$)接近0.02,两者之比是1∶200,也就是说,在10 000个个体中,只有1个个体表现隐性性状,携带这个隐性基因的杂合体却多达200个。

人为选择不能根除自然群体中有害基因的另一个缘由是,由于各种已知的和未知的原因,新的隐性突变仍将不断地在群体中产生。因此,人为限制生育的这种"选择",充其量只能是稍微降低一点突变等位基因在群体中的平衡频率,而难以实现从群体中清除隐性有害基因的目的。

前已述及,一般而言,任何一个人均可能带有4~8个隐性有害基因。因此,限制隐性突变基因携带者生育以求清除有害基因的想法也是极为荒唐的。

三、中国的"优生优育"不是Eugenics

我国从控制人口爆炸的角度出发,大力提倡计划生育与优生优育。这本是符合我国国情、利国利民的基本国策。国际上有些人却误认为我国也在搞什么"优生运动",对我国横加指责。其实我们说的优生优育(well bear and well rear)是指通过保健、咨询、教育等手段来帮助父母生出一个身心健康的孩子。并不意味着通过法律手段干涉遗传疾病患者或残疾人婚育的自主权,跟"Eugenics"的特定含义并不相干。

医学遗传学的直接目的应该是更为有效地治疗或预防遗传疾病。在与婚育有关的问题上,政府、医师及遗传学家应该通过向当事人提供遗传咨询服务等方式,帮助当事人就他们个人的婚育问题作出符合他们最佳利益的决定,从而使之及其家庭获得幸福。减少人口中残疾人的比例则只是间接目的。

Summary

Population genetics is not only an experimental but also theoretical subject. On one hand, by collecting data from reaserch, we can effectively illustrate the basic rules which continually vary generated by the population and estimate the mating type and genetic parameters of mutation or natural selection. On the other hand, according to the classical population genetics theory established by RA Fisher, S Wright and JBS Haldane in the 1930s, we can predict the constitution and variation of the population genetic, and study its rule of temporal and spatial variation of the its constitution when different factors affect it.

Genetic variation is the main focus of population genetics. But only phenotypic variation is usually observed. Sometimes the phenotypic variation is depended on the genotypic variation. Because of this role, the genotype is directly realted to the phenotype, or in other words, the phenotype is completely independent of environment. For most characters like height, weight and blood pressure, any individual's phenotype is influenced by his/her genotype and the environment (s)he lives in. At present, it is still very challenging to determine the genes affecting these polygenic characteristics. Therefore, the study of human population genetics is largely limited to the characters controlled by a single gene. In this chapter, we discuss the rule of genetic constitution and variation of qualitative characters that show a one-to-one relationship between genotype and phenotype. Then, we will discuss how genetic epidemiologic approaches, such as linkage analysis and case-control studies, can be used to assess genetic factors causing inherited diseases.

(罗泽伟)

参考文献

1. 刘祖洞. 群体遗传学//杜传书,刘祖洞. 医学遗传学. 第2版. 北京:人民卫生出版社,1992,386-412
2. Griffiths AJF, Miller JH, Suzuki DT, *et al*. An introduction to genetic analysis. Seventh edition. New York: W. H. Freemanand Company, 2000, 714-741
3. 刘祖洞. 遗传和进化//刘祖洞. 遗传学(下). 第2版. 北京:高等教育出版社,1991,318-336
4. Hartl DL, Clark AG. Principles of population genetics. Second edition. Sunderland, Massachusetts: Sinauer Associates, Inc, 1989, 118-134
5. 冯波. 群体遗传//陈仁彪,冯波. 医学遗传学. 上海:上海科学技术文献出版社,1994,143-157
6. 邱仁宗. 人类基因组研究与遗传学的历史教训. 医学与哲学,2000,21(9):1-5
7. Crow JF, Motoo Kimura. An introduction to population genetics theory. Burgess publishing company, 1970, 117

第八章 生化遗传病

生化遗传病包括先天性代谢差错（inborn error of metabolism）和分子病（molecular disease）。这个领域的创始人是19世纪末、20世纪初的英国内科医师Archibald Garrod（1858～1936），人称先天性代谢差错之父。1909年发表《先天性代谢差错》一书，系统地阐述了他对代谢病遗传基础及基因与酶关系的研究。

人类生化遗传病研究首先从血红蛋白病（hemoglobinopathy）取得突破。1949年Pauling等发现正常血红蛋白和镰状红细胞贫血症患者血红蛋白的电泳速率有异，提示这两种血红蛋白分子有化学结构上的差异，发表了《镰状红细胞贫血症，一种分子病》这一开辟新领域的重要论文。几年后，Ingram（1956）测定了镰形红细胞贫血症患者血红蛋白的氨基酸取代。由此引出40多年来对血红蛋白分子的精深研究，成为人类分子生物学和分子遗传学研究中最精彩的篇章。Gibson（1948）发现隐性遗传高铁血红蛋白血症是由于还原高铁血红蛋白的NADH黄递酶缺乏。这是人类遗传性酶病（enzymopathy）的首报病例。紧接着发现了von Gierke病（糖原累积症Ⅰ型）系由于葡萄糖-6-磷酸酶缺乏（Cori等，1952），苯丙酮尿症系由于苯丙氨酸羟化酶缺乏（Jervis，1953）。由此建立了有关人类各种遗传性

酶病的知识宝库。

生化遗传病绝大多数是常染色体隐性遗传，少部分为X伴性隐性遗传，极少部分可为常染色体显性遗传。

第一节 血红蛋白病与珠蛋白生成障碍性贫血

早在1866年人们就发现胎儿血红蛋白与成人血红蛋白有所不同。1925年Cooley描述了一个幼儿患有贫血，并伴发有核红细胞增多症。此后意大利、希腊等地中海区域也陆续发现并报道了类似疾病，后来被称作珠蛋白生成障碍性贫血，因这类疾病首先在地中海区域被发现，因此当时又称为地中海贫血(thalassemia)。另一方面，1904年在一黑人血液中发现了镰形红细胞。1910年报道一例居住在芝加哥的西印度群岛青年男子患有严重镰形红细胞贫血。1917年发现此类贫血患者红细胞可在体外发生镰变(sickling)。1927年发现缺氧是这种红细胞发生镰变的重要条件，在氧气充足时可以恢复原状。前已述及，Pauling等(1949)发现镰形红细胞的血红蛋白有不同于正常血红蛋白的电泳性质，认为必定是两者珠蛋白分子有所不同。几年后，In-gram(1956)鉴定了这两种血红蛋白的不同是在于β链第6位氨基酸，正常者为谷氨酸，而在镰形红细胞贫血症中则是缬氨酸。此后各种异常血红蛋白被陆续检出，至今全世界已发现异常血红蛋白1000多种，据不完全统计，全世界异常血红蛋白携带者有1亿多人。

一、血红蛋白的分子结构及其遗传控制

血红蛋白是一种含有色素辅基的结合蛋白，其色素部分是血红素(heme，haem)，蛋白质部分是珠蛋白(globin)，血红蛋白是一个四聚体，它是由两条α链和两条非α链组成。这些肽链被两个不同的基因簇所编码。α基因簇在第16号染色体的短臂16pter-p13.3(图8-1)，在α基因簇中有两个表达的α珠蛋白结构基因(α_1和α_2)，再有在胚胎期表达的ξ2基因和ψξ1、ψα1、ψα2和θ等4个假基因(pseudo-gene)，假基因是一种与结构基因具有同源顺序的基因，但是它已发生了突变而失去了表达作用。β基因簇在第11号染色体短臂11p15.5(图8-2)，依次排列的顺序是胚胎期的ε珠蛋白基因、胎儿$^{G}\gamma$和$^{A}\gamma$珠蛋白基因、假基因ψβ、δ和β(成人)珠蛋白基因。在16号染色体短臂α基因簇5′端上游60kb处，有α位点控制区(α-locus control region，α-LCR)，它对α珠蛋白基因簇结构基因的表达有十分重要的调控作用。在11号染色体短臂上也有类似的发现，称为β位点控制区(β-locus control region，β-LCR)。

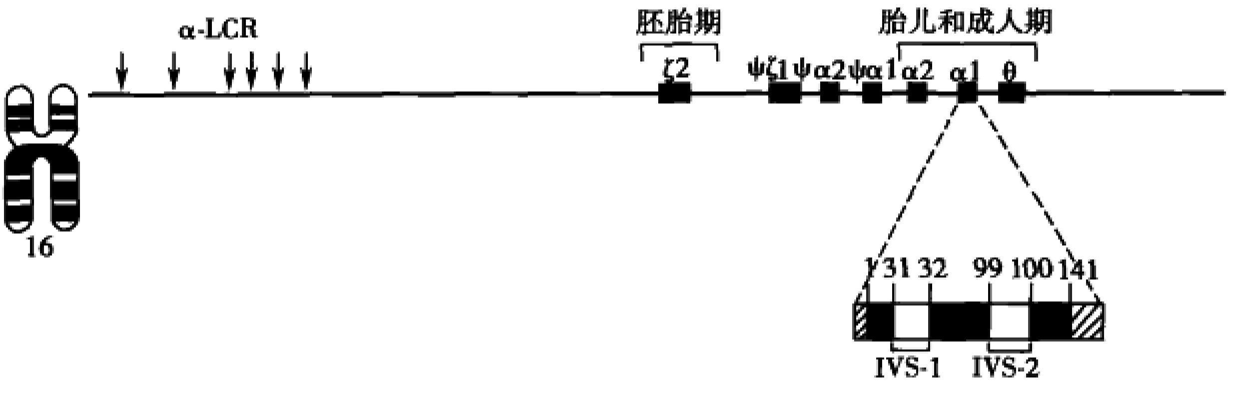


图8-1 α-珠蛋白基因簇和α-珠蛋白基因的结构

ψξ1、ψα2、ψα1、θ：假基因 α-LCR：α位点控制区

每一个能表达的珠蛋白基因有三段外显子和两段间隔顺序(内含子)。转录形成的mRNA前体在加工过程中切除间隔顺序，并把外显子连接在一起，构成一条成熟的mRNA分子，由此翻译成α珠蛋白链或非α珠蛋白链。α珠蛋白链含有141个氨基酸，非α珠蛋白链有146个氨基酸。血

红蛋白分子由两条 α 链和两条非 α 链组成。正常人体从胚胎到成人，可以有六种血红蛋白类型（图 8-3）。

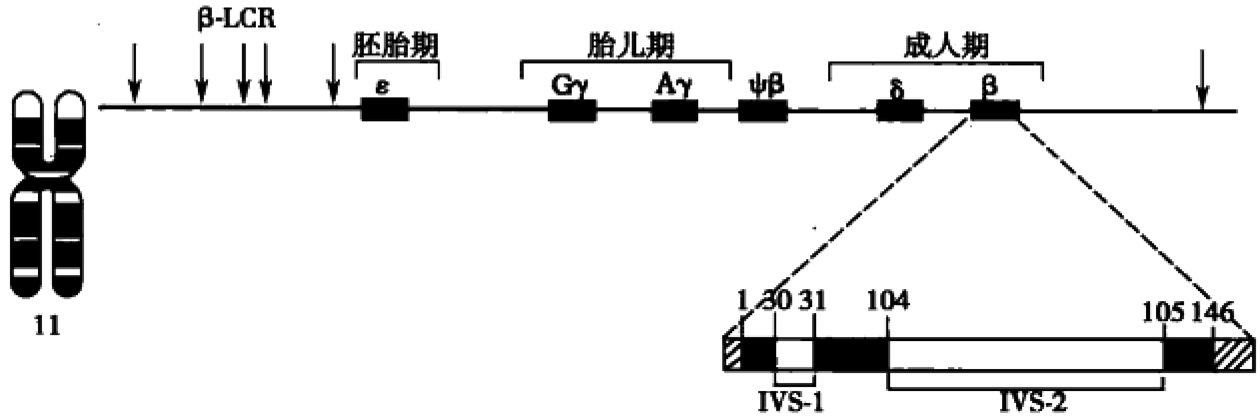


图 8-2 β-珠蛋白基因簇和 β-珠蛋白基因的结构

ψβ：假基因 β-LCR：β 位点控制区

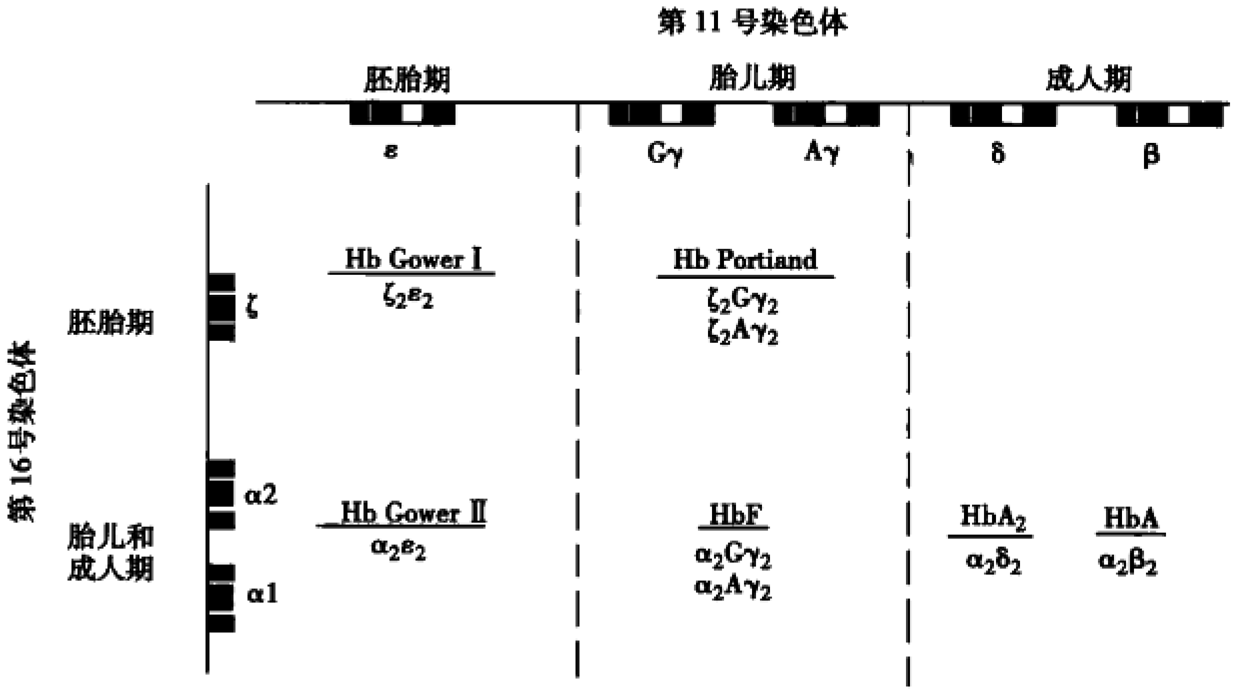


图 8-3 正常人体发育过程中的血红蛋白分子类型

极少数异常血红蛋白分子由 4 条同样肽链组成，如 Hb H 为 β_4，Hb Bart's 为 γ_4。在人体发育的不同阶段，各种血红蛋白的合成呈现严格的消长过程（图 8-4）。Hb Gower 2 和 Hb Gower 1 仅见于胚胎发育早期，持续到约第 8 周，大概是原始卵黄囊红细胞发生系统的产物。Hb Portland 也仅见于胚胎期。ξ 链在 Hb Bart's 胎儿水肿综合征死胎中有相当数量，在正常脐血中仅含微量。ε 链合成的时间

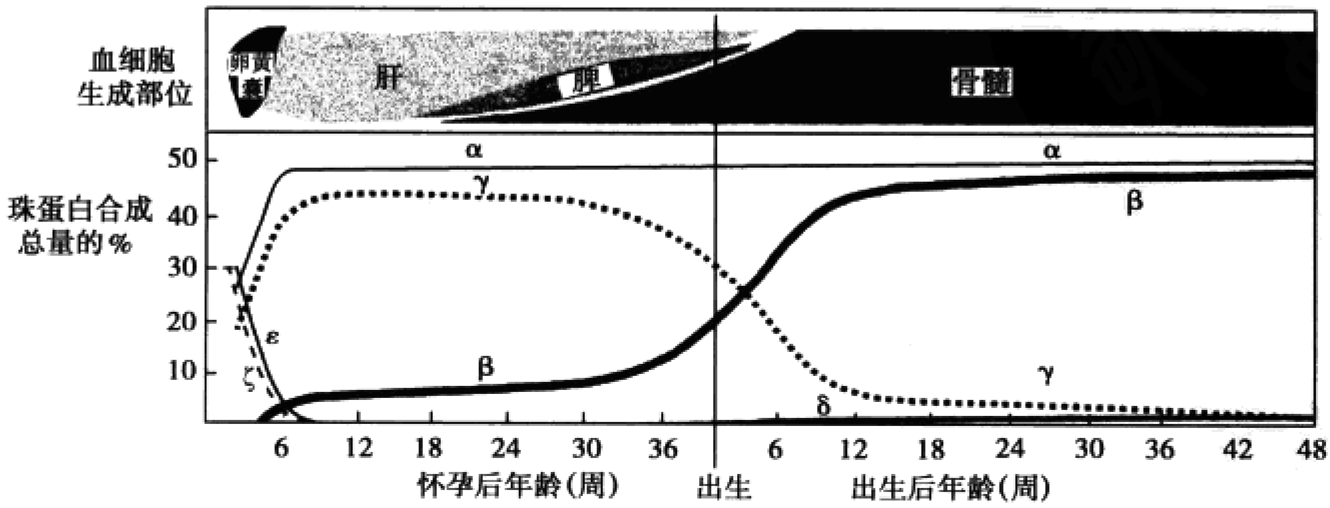


图 8-4 正常人体发育过程中珠蛋白肽链合成的演变

亦在胚胎早期，继后就开始合成 α 链和 γ 链。α 链首先与 ε 链合成 Hb Gower 2，然后与 γ 链合成 Hb F。Hb F 是胎儿主要血红蛋白。胚胎发育至第 8 周时开始合成 β 链，此时仅占 β 链簇的 10% 左右，此后 β 链合成率迅速增高，而 γ 链合成率迅速降低。出生后不久，合成大致等量的 γ 链和 β 链，然后 β 链迅速增高，而 γ 链却减少，成人中仅有少量 γ 链构成微量 Hb F。γ 链中 $^{G}\gamma$: $^{A}\gamma$ 在出生时为 3:1，而在正常人 Hb F 中成为 2:3。δ 链合成的确切时间还不清楚，可能在出生前几周内开始合成。脐血中有微量 δ 链，1 周岁时已达到成人水平。在正常发育过程中，胚胎期的血红蛋白是 Hb Gower 1，Hb Gower 2 和 Hb Portland。胎儿期主要是 Hb F。成人有三种血红蛋白，Hb A 占总量 98%，Hb A_2 约占 2%，另外还有微量 Hb F。

二、血红蛋白变异体与血红蛋白病

（一）血红蛋白变异体

自从 1949 年 Pauling 发现镰形细胞血红蛋白 Hb S 以来，已由溶血性贫血、红细胞增多或青紫患者及群体电泳筛查发现 1000 多种遗传性血红蛋白结构异常。这些血红蛋白变异体多半有电荷的改变，因而有不同的电泳迁移速度。图 8-5 是 Hb A、Hb S 和 Hb C 的电泳结果。血红蛋白的分子结构可以缩写表示，如将 Hb A 写成 $\alpha_2^A\beta_2^A$，将 Hb F 写成 $\alpha_2^A\gamma_2^F$。在异常血红蛋白，则写出其异常肽链名称，并标明取代氨基酸的位置，如 Hb S 可写成 $\alpha_2^A\beta_2^{6缬}$ 或 $\alpha_2\beta_2^{6谷\to缬}$。

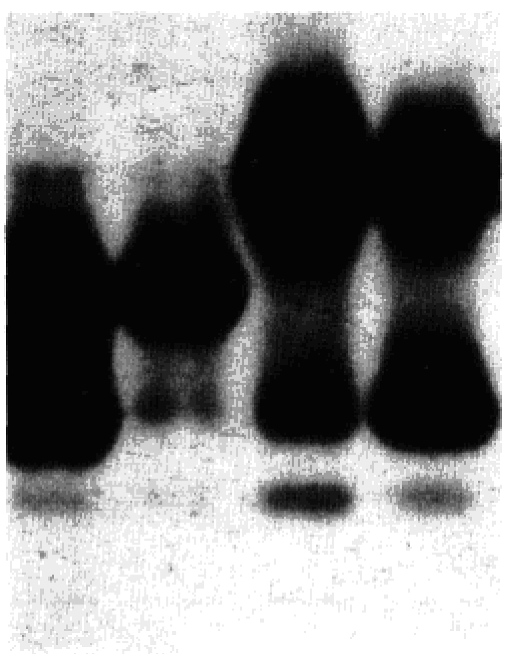


图 8-5 Hb A、Hb S 和 Hb C 的电泳结果
1. Hb S/Hb C 2. Hb S
3. Hb A 4. Hb A/Hb C

1. 单个碱基置换 大多数的血红蛋白结构异常是由于肽链中单个氨基酸残基为另一个氨基酸残基所取代，其原因是相应的遗传密码发生了单个碱基转换或颠换，如 Hb $G_{Chinese}$（$\alpha^{30谷\to谷胺}$）是由于 G→C，Hb S（$\beta^{6谷\to缬}$）是由于 A→U。

在珠蛋白肽链基因的终止密码（UAA、UAG 或 UGA）发生碱基置换而成为可读密码时，则肽链将延长直至下一个终止密码。如正常 α 链有 141 个氨基酸残基，而 Hb Constant Spring 的 α 链延长到 172 个残基。那是因为第 142 个密码子 UAA（终止）→CAA（谷胺），肽链继续合成到第 173 处 UAA（终止）之前。

2. 密码子缺失或插入 Hb Lyon 是由于 β 链第 17 ~ 18 位决定赖氨酸和缬氨酸的两个密码子的缺失，从而造成两个氨基酸的丢失，第 16 位及其以前的氨基酸顺序与第 19 位及其以后的氨基酸顺序均无改变。Hb Grady 则是由于 α 链第 116 ~ 118 位插入了谷-苯丙-苏三个氨基酸残基，显然是由于在 α 珠蛋白基因中插入了为这三个氨基酸编码的 9 个碱基。

3. 移码突变 在珠蛋白肽链基因的碱基缺失或插入不是 3 个或 3 的倍数个碱基时，这将导致突变部位以后的密码子变化，重新组合三联密码子而导致移码（frame-shift）。如 Hb tak 是 β 链在第 146 ~ 147 个氨基酸密码子（CAC ~ UAA）之间插入了 AC，使终止密码 UAA 变成苏氨酸的密码 ACU，肽链延长到 157 个氨基酸。

4. 融合基因 两种非同源基因的部分片段拼接而成的基因，称为融合基因（fusion gene）。这种融合基因可能是在减数分裂时同源染色体之间错位配对引发不等交换的结果。Hb Lepore 的 α 链氨基酸顺序正常，其非 α 链由 δβ 融合基因编码，肽链 N 端为不等交换产生的 δ 链氨基酸顺序，C 端为 β 链氨基酸顺序，构成 δβ 链。对应融合基因为 βδ 融合基因，见于 Hb Anti-Lepore（图 8-6）。

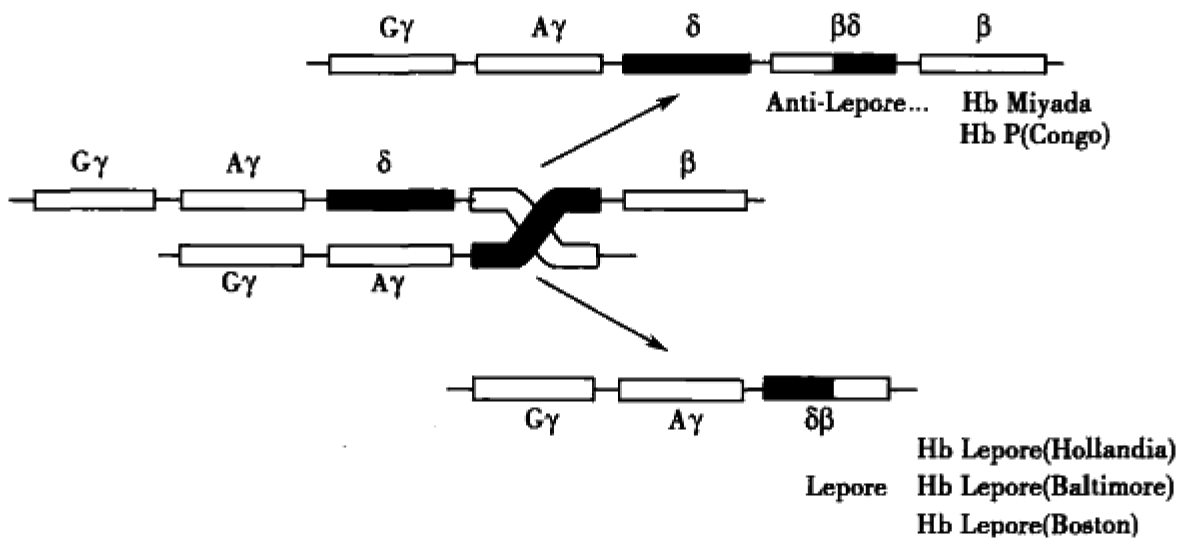


图 8-6 δ 珠蛋白基因和 β 珠蛋白基因发生不等交换导致 δβ 和 βδ 融合基因

(二) 血红蛋白病

大多数血红蛋白变异体是稀有的,半数以上并不致病,在正常人群中存在着变异体,这只有在进行群体筛查时才能发现。近年来,我国在血红蛋白变异体的研究方面取得了不少成绩,亦发现了不少新的变异体,如 Hb Harbin。有些血红蛋白变异体引起功能上的改变,如溶解度降低(如 Hb S),或是导致分子不稳定(如 Hb Zurich),或是对氧的亲和力增高(如 Hb Chesapeake)或降低(如 Hb Kansas),或是形成不能带氧的高铁血红蛋白(methemoglobin)(如 Hb Boston),从而产生血红蛋白病。临床最主要的症状是溶血性贫血(hemolytic anemia),红细胞代偿性增多和青紫(cyanosis)。

1. 镰状红细胞贫血症(sickle cell anemia)　本症为常染色体隐性遗传,是由于 β 链第 6 位谷氨酸被缬氨酸取代,成为 Hb S。纯合子患者($\alpha\alpha\beta^S\beta^S$)血红蛋白溶解度下降,使红细胞镰变(图 8-7)。纯合子症状严重,可产生血管阻塞危象,阻塞部位不同可引起不同部位的异常反应如腹部疼痛、脑血栓等,另有严重溶血性贫血及脾大等症状。杂合子个体 $\alpha\alpha\beta^A\beta^S$ 不表现临床症状,但是在氧分压低的情况下可以引起红细胞镰变,称为镰状红细胞性状(sickle cell trait)。

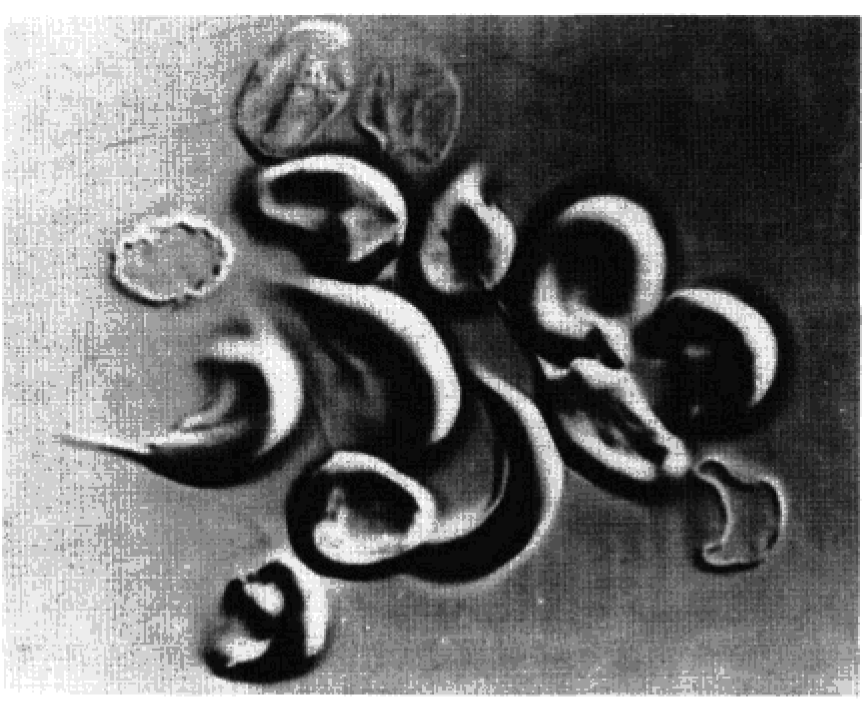


图 8-7　Hb S 纯合子的镰形红细胞

以 β 珠蛋白基因有关顺序为探针作分子杂交,可由 RFLP 对镰状红细胞贫血症作分子遗传诊断。内切酶 Mst Ⅱ 识别 CCTNAGG,正常 β 珠蛋白基因(CCT-GAG-GAG)产生 1.15kb 和 0.20kb 的片段。而 Hb S 的 β 珠蛋白基因(CCT-GTG-GAG)失去了这个 Mst Ⅱ 识别顺序,产生 1.35kb 片段(图 8-8)。

2. Hb C 病　本症亦为常染色体隐性遗传,是由于 β 链第 6 位谷氨酸被赖氨酸取代。纯合子患者($\alpha\alpha\beta^C\beta^C$)呈现轻度溶血性贫血,可伴有肝脾肿大。杂合子($\alpha\alpha\beta^A\beta^C$)可无临床症状。

Hb C($\beta^{6谷\rightarrow赖}$)与 Hb S($\beta^{6谷\rightarrow缬}$)属等位变异,都是 β 珠蛋白链第 6 位的氨基酸取代,这两个珠蛋白基因可称为真等位基因(eualleles)。同一肽链但不是同一位置氨基酸取代则属异等位基因(heteroalleles),如 Hb C($\beta^{6谷\rightarrow缬}$)与 Hb $M_{Saskatoon}$($\beta^{63组\rightarrow酪}$)的 β 珠蛋白基因。

3. Hb M 遗传性高铁血红蛋白血症(methemoglobinemia)　本症为常染色体显性遗传,是由于取代的氨基酸,如 Hb M_{Boston}($\alpha^{58组\rightarrow酪}$),α 链 58 位取代了组氨酸的酪氨酸占据了血红素 Fe 原子的配基

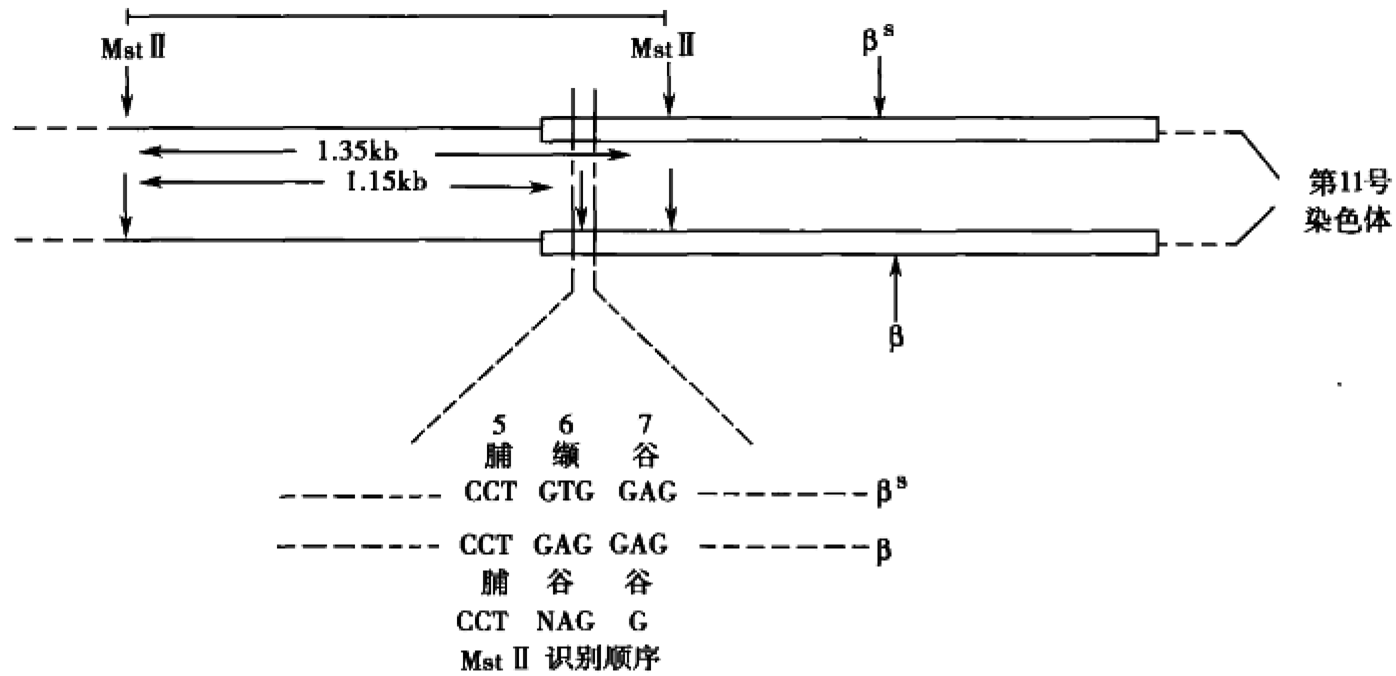


图 8-8 镰形红细胞贫血症的分子诊断

位置,使 Fe 原子呈稳定高铁状态,丧失了血红素与氧结合的能力,导致组织缺氧。患者呈现发绀症状并导致继发性红细胞增多(erythrocytosis)。已知的高铁血红蛋白尚有 Hb M_{Iwate}($\alpha^{87组\rightarrow酪}$)、Hb $M_{Sakatoon}$($\beta^{58组\rightarrow酪}$)、Hb $M_{Hydepark}$($\beta^{92组\rightarrow酪}$)和 Hb $M_{Milwaukee}$($\beta^{63缬\rightarrow谷}$)。

4. Hb Bristol 不稳定血红蛋白病　本症亦为常染色体显性遗传,Steadman 等人在 1970 年报道他们的研究:这种疾病的遗传机制是 β 链第67 位缬氨酸被天冬氨酸所取代,导致血红蛋白分子不稳定。这种不稳定的血红蛋白易在细胞中发生变性沉淀而形成 Heinz 小体,因而造成溶血性贫血。Ree 等人在 1996 年报道用 DNA 和蛋白质分析发现这一定位是错的。他们重新研究了几位被报道的自发性 Heinz 小体溶血性贫血,发现在 RNA 翻译成蛋白质后,67 位的天冬氨酸是由原先的蛋氨酸经过缓慢地静止性(silent)修饰而来,也就是说,DNA 上发生的突变实际是 67 位缬→蛋(甲硫)。这种转录翻译后的修饰可能是由邻近的血红素和氧分子催化。这一发现说明同时分析蛋白质和 DNA 在研究中的重要性。这种类似的情况还发生在 Hb Coventry。本症主要临床症状是先天性溶血性贫血,黄疸和脾肿大,因而亦称先天性 Heinz 小体溶血性贫血(congenital Heinz body hemolytic anemia)。已知的不稳定血红蛋白有 90 余种。

Hardison 等人从 2002 年起创建了 HbVar(http://globin.cse.psu.edu/globin/hbvar/),可以在网上查阅珠蛋白变异的所有新老信息。目前,HbVar 还可以和 GALA[基因组联盟和注释信息 Genome Alignment and Annotation database(http://gala.cse.psu.edu)]联网,可以从中获取更多的相关信息。

三、珠蛋白生成障碍性贫血

珠蛋白生成障碍性贫血是由于珠蛋白多肽链完全不能合成或合成不足所致。成人血红蛋白(Hb A)由两条 α 链和两条 β 链组成。珠蛋白生成障碍性贫血中凡由于 α 链合成减少或不能合成者称为 α 珠蛋白生成障碍性贫血,由于 β 链合成减少或不能合成者称为 β 珠蛋白生成障碍性贫血。

α 或 β 珠蛋白链合成减少或甚至完全缺如涉及 α 或 β 珠蛋白基因的种种突变,可分为非缺失型(包括微缺失型)和缺失型两大类。非缺失型珠蛋白生成障碍性贫血涉及从 5′转录控制信号、外显子密码、内含子(间隔顺序)拼接信号和共有序列(consensus sequence)、外显子和内含子潜在的拼接部位、终止密码和 3′多聚腺苷化信号等处的碱基取代、缺失、插入、移码等,导致转录受阻或转录产物异常,使 RNA 加工拼接或翻译受阻,RNA 不稳定或翻译后异常肽链不稳定,最终导致患者 α 或 β 珠蛋白链减少(α^+ 或 β^+ 珠蛋白生成障碍性贫血)或完全缺如(α^0 或 β^0 珠蛋白生成障碍性贫血)。缺失型珠蛋白生成障碍性贫血涉及 α 或 β 珠蛋白基因簇较大范围的缺失,包括涉及 α 或 β 珠蛋白基因簇 5′

上游60kb处的α或β位点控制区缺失。大部分全缺失发生在α或β位点控制区。

大多数α珠蛋白生成障碍性贫血基因是缺失一个α珠蛋白基因，是不等交换的结果。这是因为串联重复基因(tandemly repeated genes)易于在减数分裂时发生非同源配对而发生不等交换。在1条染色体上只有1份的基因，其缺失是少见的，只占突变的大约1% ~5%。β珠蛋白生成障碍性贫血基因大约只有1%是由于缺失或重排。

(一) α珠蛋白生成障碍性贫血

引起α珠蛋白生成障碍性贫血的α珠蛋白基因主要突变类型和主要缺失类型分别见表8-1和图8-9。

表8-1 α珠蛋白生成障碍性贫血的α珠蛋白基因的突变类型

突变类型	分子缺陷	来源
1. 核苷酸取代		
(1) α_2 起始密码子 ATG→ACG	无功能mRNA	地中海地区人
(2) Hb Quong Sze ($\alpha^{125亮\to脯}$)	α-β二聚体形成受阻	中国人
(3) α_2 多聚A信号 AATAAA→AATAAG	RNA切割和多聚腺苷化突变体	沙特阿拉伯人
(4) Hb Constant Spring ($\alpha^{终止\to谷胺}$)	终止密码突变，肽链延长，但mRNA极不稳定	东南亚人
2. 移码突变		
α_1 密码子14TGG-T	无功能mRNA	沙特阿拉伯人
3. 核苷酸缺失		
α_2IVS-1共有顺序缺失5个核苷酸	RNA拼接缺陷	地中海地区人

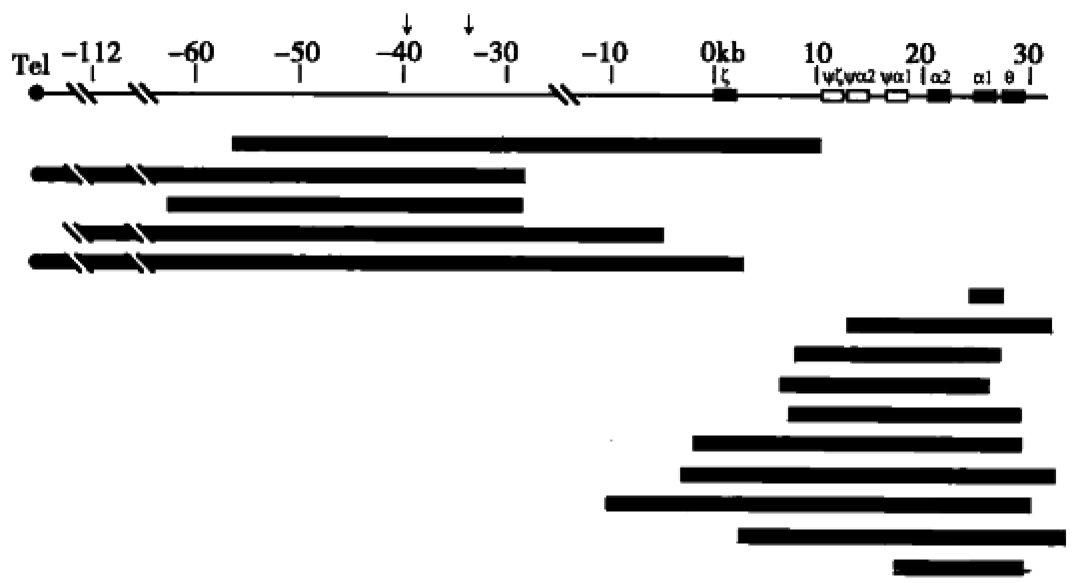


图8-9 α珠蛋白基因簇主要缺失类型

箭头示α位点控制区

人体第16号染色体短臂上有2个连锁的α珠蛋白基因。如果16号染色体上的2个α基因均因突变或缺失而丧失了功能，我们把这种单倍型称为α地$_1$(α-thal$_1$)。如果16号染色体上的2个α基因只有1个因突变或缺失而丧失了功能，这种单倍型称为α地$_2$(α-thal$_2$)。α珠蛋白生成障碍性贫血

杂合子也有一定的临床表现，故本症属常染色体显性遗传。因为患者丧失功能的 α 基因的数目不同，α 珠蛋白生成障碍性贫血有以下四种临床类型（图 8-10）。

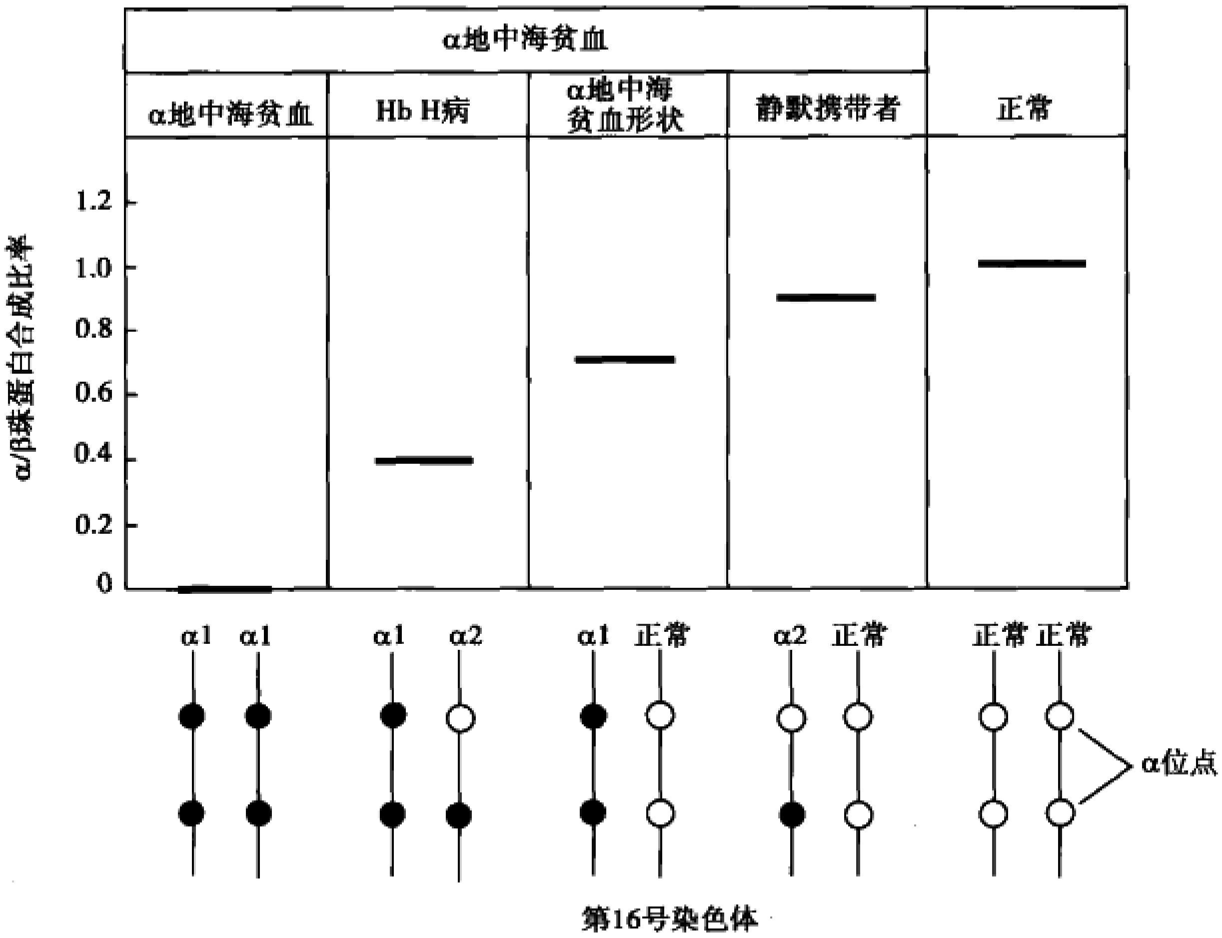


图 8-10 不同临床类型 α 地中海贫血的基因型与 α/β 珠蛋白合成比率

1. Hb Bart's 胎儿水肿综合征（hydrops fetalis syndrome） 本症为 α^0 珠蛋白生成障碍性贫血。胎儿 4 个 α 基因均丧失功能（－－/－－），无 α 珠蛋白链合成，80% 以上血红蛋白为 Hb Bart's（$^A\gamma_4$、$^G\gamma_4$），其余为 Hb H（β_4）和 Hb Portland（$\zeta_2{}^A\gamma_2$、$\zeta_2{}^G\gamma_2$）。Hb Bart's 有很高的氧亲和力，致使组织严重缺氧。胎儿严重水肿，导致自发性流产或出生后不久死于严重水肿。本症患儿为常染色体显性纯合子，父母均为 α 地$_1$ 单倍型杂合子。

2. Hb H 病 患者有 3 个 α 基因丧失功能（－－/α－），亲代分别为 α 地$_1$（－－）和 α 地$_2$（α－）携带者。由于 α 珠蛋白链严重短缺，胎儿出生时 Hb Bart's（$^A\gamma_4$、$^G\gamma_4$）达 20% 或更高，在从 γ 链到 β 链的转换完成后 Hb H（β_4）达 4% ~30%。本症属中度珠蛋白生成障碍性贫血（thalassemia intermedia），Hb Bart's 和 Hb H 沉淀形成包涵体（in-clusion body），导致低色素小细胞性溶血性贫血，患者有黄疸和肝脾肿大。

3. α 珠蛋白生成障碍性贫血性状 患者有 2 个 α 基因丧失功能，在东方人通常为 α 地$_1$ 杂合子（－－/αα），而在黑人中通常为 α 地$_2$ 纯合子（α－/α－）。本症为轻度 α 珠蛋白生成障碍性贫血（tha-lassemia minor），患者有轻度小细胞性贫血。

4. α 珠蛋白生成障碍性贫血静止型携带者 受累者仅有 1 个 α 基因丧失功能（αα/α－），有正常血象，可无临床症状。

Hb H 病、α 珠蛋白生成障碍性贫血性状和静止型携带者均有不同程度的 α 珠蛋白链合成，可合称为 α^+ 珠蛋白生成障碍性贫血，患者均属杂合子。

（二）β 珠蛋白生成障碍性贫血

引起 β 珠蛋白生成障碍性贫血的 β 珠蛋白基因主要突变类型和主要缺失类型分别见表 8-2 和图 8-11。本症为常染色体隐性遗传，纯合子呈现重型珠蛋白生成障碍性贫血，杂合子表现为 β 珠蛋白生成障碍性贫血性状。

表8-2 β珠蛋白生成障碍性贫血的β珠蛋白基因的突变类型

突变类型	分子缺陷	来源
1. 核苷酸取代		
(1) β-28 核苷酸 A→C	TATA 区转录信号突变, β^+	中国人
(2) βIVS-1 第1核苷酸 G→A	拼接接头突变，导致 RNA 加工缺陷, β^0	地中海地区人
(3) βIVS-1 第5核苷酸 G→C	共有顺序突变导致 RNA 加工缺陷, β^+	印第安人
(4) βIVS-2 第654核苷酸 C→T	内含子变化导致 RNA 加工缺陷, β^0	中国人
(5) β第17密码子 A→T	无义突变导致无功能 RNA, β^0	中国人
(6) Hb Knossos($\beta^{27丙\to丝}$)	编码区突变导致 RNA 加工缺陷，静止型 β^+	地中海地区人
(7) β多聚A信号	RNA 切割和多聚A化信号突变, β^+	美国黑人
AATAAA→AACAAA		
2. 核苷酸插入		
β密码子 71~72 间插入	移码突变导致无功能 mRNA, β^0	中国人
1个核苷酸		
3. 核苷酸缺失		
βIVS-1 缺失 25bp	拼接接头改变导致 RNA 加工缺陷, β^0	印度人
4. 不等交换		
Hb Lepore	δβ 联接 mRNA，不稳定, β^0	不明

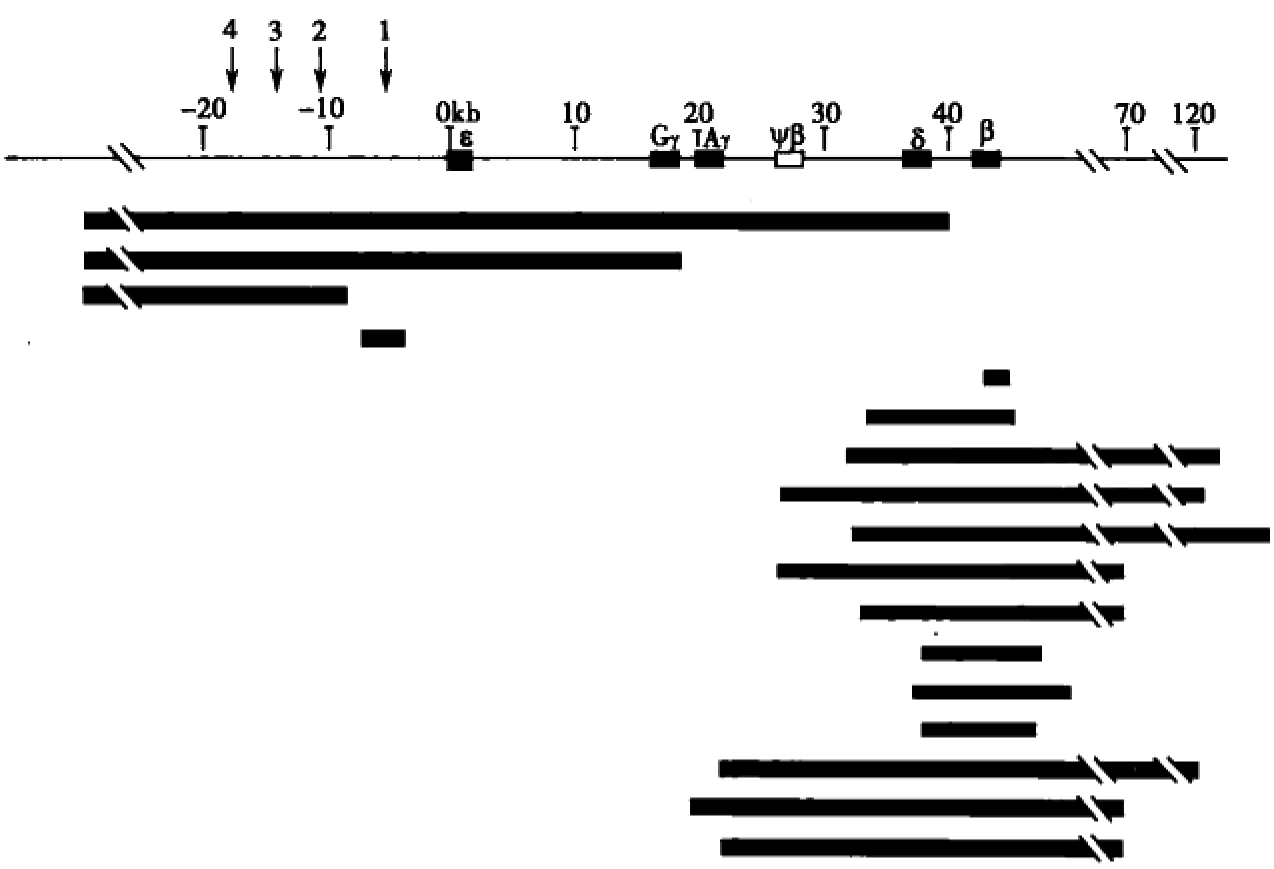


图8-11 β珠蛋白基因簇主要缺失类型

箭头示β位点控制区

不同人种群最常见的β珠蛋白生成障碍性贫血突变见表8-3。

表8-3 不同人种群最常见的β珠蛋白生成障碍性贫血突变

人种群	β珠蛋白生成障碍性贫血突变	频率
1. 中国人	β密码子71/72 +1移码	49%
	βIVS-2第654核苷酸C→T	38%
2. 地中海人	βIVS-1第110核苷酸G→A	35%
	β密码子39无义突变C→T	27%
3. 美国黑人	TATA区-29A→G	39%
	多聚部位T→C	26%
4. 印度人	βIVS-1核苷酸G→C	36%
	部分缺失(619核苷酸)	36%

1. β^0珠蛋白生成障碍性贫血 β^0珠蛋白生成障碍性贫血亦称重型珠蛋白生成障碍性贫血(thalassemia major)或Cooley贫血。患者为β地(β-thal)基因纯合子(β地/β地)。患婴出生时正常，半周岁时发生严重小细胞性溶血性贫血，过剩的α珠蛋白沉淀为包涵体。患儿因骨髓外血细胞生成而发生肝脾肿大和β珠蛋白生成障碍性贫血面容，呈大头、颧骨突出、塌鼻梁、眼距过宽和脸水肿。患者血红蛋白水平低于5g/dl，有人统计英国50%的此类患者在35岁前死亡，死因往往是传统输血治疗引起的铁负荷过重。但现在可以通过一些措施延长患者寿命，包括：①通过植入装置静脉输入去铁胺(desferrioxamine)24小时去铁，对铁负荷严重超标的患者有效；②每周家庭使用一次性输液器输入去铁胺(如Baxter气球泵)；③口服铁螯合剂deferiprone(L1)。当β地(β-thal)基因纯合子患者同时合并有α珠蛋白生成障碍性贫血，四个α珠蛋白基因中仅有三个为有功能的基因时，症状往往较轻，不需要常规输血治疗，这种情况被称为β珠蛋白生成障碍性贫血中间型(β thalassemia intermedia)。人们推测可能是多余α珠蛋白链的量减少之后使得体内无效造血情况减轻，临床症状也随之减轻。同样人们也观察到当体内γ珠蛋白合成增加时也能减轻患者的贫血症状。

2. β^+珠蛋白生成障碍性贫血 β^+珠蛋白生成障碍性贫血即β珠蛋白生成障碍性贫血性状，受累者为β地杂合子(β地/β)，通常仅有轻度贫血，但无明显临床症状。约半数病例有轻度到中度脾肿大。α珠蛋白表达情况对β珠蛋白生成障碍性贫血的影响在这种疾病中也有例可查。人们发现有些患者为同一条16号染色体上有三个功能性α珠蛋白基因(α珠蛋白过表达)的β地杂合子，症状明显加重，和重型珠蛋白生成障碍性贫血的患者相同，他们也需要常规输血治疗以及脾切除术。有人推测这可能是由于患者无法水解多余的α珠蛋白链所引起。

第二节 血浆蛋白病

血友病

(一) 血友病A

血友病A(hemophilia A)是由于Ⅷ凝血因子即抗血友病球蛋白(antihemophilic globulin，AHG)遗传性缺乏。Ⅷ因子有两个成分。Ⅷc(Ⅷ因子凝血成分)具有Ⅷ因子凝血活性，$Ⅷ_{Ag}$(Ⅷ因子相关抗原)是Ⅷ因子凝血活性的载体蛋白，仅能由免疫学方法检出。血友病A患者的Ⅷc含量低或无，但有正常或高的$Ⅷ_{Ag}$。

血友病的主要临床表现是反复自发性或在轻微损伤后出血不止。体表、体内任何部分均可出血，可以涉及皮肤、黏膜、肌肉内或器官内，如关节腔出血可致关节积血。

本症为X连锁隐性遗传。*AHG*基因位于Xq28的近侧，长186kb，有26个外显子，编码AHG的

2 332个氨基酸。基因突变涉及核苷酸取代、缺失、插入和移码，其中半数以上为各种程度缺失见表8-4。

表8-4 Ⅷ因子突变类型与血友病A临床严重程度

Ⅷ因子分子缺陷	血友病A临床	Ⅷ因子分子缺陷	血友病A临床
1. 核苷酸取代		(12)精2307亮(G→T)	中度
(1)谷272甘*(A→G)	中度	(13)精2307终止(C→T)	重度
(2)精372组(G→A)		2. 缺失	
(3)酪1680苯丙(A→T)	轻度	(1)IVS-1缺失7kb	正常
(4)谷1686终止(C→T)	重度	(2)缺失外显子1	重度
(5)精1689半胱(C→T)	中度	(3)缺失外显子1~26	重度
(6)酪1709半胱(A→G)		3. 插入	
(7)精1941谷胺(G→A)	轻度	(1)IVS-10插入0.7kb	正常
(8)精1941终止(C→T)	重度	(2)外显子14插入3.5kb	重度
(9)精2116终止(C→T)	重度	4. 移码	
(10)精2209谷胺(G→A)	重度	(1)外显子8密码子360GAA缺失GA	重度
(11)精2209终止(C→T)	重度		

*表示第272位谷氨酸突变为甘氨酸，余类同

近年来又发现40%血友病A患者的基因缺陷是由于倒位引起的。上海瑞金医院血液学研究所张宇舟等应用Southern印迹分析检测66例血友病A患者，其中22例X染色体长臂内的FⅧ基因内含子22部位有倒位现象，且这些患者的临床症状都很严重，占重型血友病A患者的46.6%。这对于明确血友病A的发病机制有着相当重要的意义。

治疗主要是输入人血浆中提炼或通过重组技术合成的Ⅷ因子，这种替代疗法在大部分病例中都有效，但有10%~15%的患者会产生中和性抗体，从而降低了治疗效果。Ⅷ因子的半衰期为8~12小时，部分患者需要每天接受两次治疗。

本症在主要人群中累及1/10 000~1/5000的男性。但也有女性患者的例子出现。Pola等人在1957年报道了一位纯合子女性患者，她的父亲是一名血友病患者，其父母亲是双一级表亲。女性患者的症状以及病情程度同男性半合子患者相仿，有时杂合子携带者发生X染色体和常染色体易位而破坏了唯一有功能的*AHG*基因时，也会发生血友病。另外一种罕见的情况是女性X染色体失活时带有倾向性，而不是随机失活，也会出现女性血友病。例如Coleman等在1993年曾报道一位女性血友病患者，母亲为失禁性色素沉着(incontinentia pigmenti，IP)患者，IP的基因定位也在X染色体，父亲为血友病患者。这位女性患者(是婴儿期)的体细胞表现出明显倾向性的X染色体失活现象：带有IP基因的X染色体更倾向于失活，所以这名女婴成为Ⅷ因子突变的功能性杂合子，表现出血友病的症状。

与血友病A有关的是von Willebrand病。患者$Ⅷ_c$和$Ⅷ_{Ag}$均减少，反复出血多累及黏膜如鼻黏膜和胃肠道，出血的严重程度可有很大的变异。von Willebrand病呈常染色体显性遗传。

von Willebrand最早在1931年于Aland岛发现该岛有一些易出血的人群，并将其命名为假血友病。与血友病A的主要区别是出血时间更长，随着年龄的增长，度过青春期后症状会有所减轻。患者$Ⅷ_c$和$Ⅷ_{Ag}$均减少，反复出血多累及黏膜如鼻黏膜和胃肠道，出血的严重程度可有很大的变异。von Willebrand病多呈常染色体显性遗传。基因定位于12p13.3。每100万人中有30~100人发病。根据临床症状的差别，可以分为三种亚型：

Ⅰ型（占了总数的60%～80%）：轻到中度缺乏VWF和Ⅷ因子，大约降到血浆正常水平的5%～30%。这种亚型多为常染色体显性遗传。

Ⅱ型（占10%～30%）：是由于VWF的功能性质（qualitative）异常所引起，还可以进一步分成2A、2B、2M和2N四种亚型，通常为常染色体显性遗传，也有少量常染色体隐性遗传的例子。

Ⅲ型（占1%～5%）：常染色体隐性遗传，杂合子或纯合子具有低于正常或无法检测到的血浆VWF（小于正常血浆水平的1%）、低水平的Ⅷ因子（1%～10%正常血浆水平）。

（二）血友病B

Biggs等人在1952年12月27日（圣诞节期间）报道了一名5岁男孩，患有类似血友病的出血症状，但是由于不同于Ⅷ因子的凝血因子缺乏所引起的。这名患者的名字为Stephen Christmas。作者当时就把这种新发现的疾病命名为Christmas病。

血友病B（Christmas病）是由于第Ⅸ凝血因子，即血浆凝血活酶成分（plasma thromboplastin component，PTC）遗传性缺乏。其主要临床症状与血友病A完全相同。

本症亦为X连锁隐性遗传。*PTC*定位于Xq27.1-q27.2，长34kb，有8个外显子，编码415个氨基酸。*PTC*基因突变涉及核苷酸取代、缺失、插入和移码，其中大部分为核苷酸取代（表8-5）。

表8-5 Ⅺ因子突变类型与血友病B临床严重程度

Ⅺ因子变异体	分子缺陷	血友病B临床
1. 核苷酸取代		
（1）F9（Seattle-3）	谷27赖*（G→A）	严重
（2）F9（Durham）	甘60丝（G→A）	严重
（3）F9（Seattle-4）	精248谷胺（G→A）	严重
（4）F9（Chongqing）	谷27缬（A→T）	严重
（5）F9（Alabama）	天冬47甘（A→G）	轻度
（6）F9（Cardiff）	精145半胱（G→T）	中度～严重
（7）F9（Hilo）	精180谷胺（G→A）	严重
（8）F9（London）	精333谷胺（G→A）	中度
（9）F9（Bonn-1）	精390终止（C→T）	严重
（10）F9（Niigata）	丙390缬（C→T）	中度～严重
（11）F9（Vancouver）	异亮397苏（T→C）	中度～严重
（12）F9	5'IVS-3拼接供者部位（GT→GG）	严重
（13）F9	3'IVS-7拼接受者部位（G→A）	严重
2. 缺失		
（1）F9	缺失外显子1～3	严重
（2）F9	缺失外显子2～8	严重
3. 插入		
（1）F9（El Salvador）	外显子4大片段插入	中度～严重
（2）F9	外显子5插入Alu顺序	严重
4. 移码		
（1）F9（Seattle-2）	天冬85缬，86终止（外显子缺失A）	严重

*表示27为谷氨酸突变为赖氨酸，余类同

对血友病B的治疗一般为输入Ⅸ因子，有些学者进行了重组腺相关病毒作为载体的基因治疗人体试验，通过肌肉注射病毒载体的方法以期达到治疗目的。虽然在人们普遍担心的安全性上没有发现严重问题，能够略微减轻疾病症状，减少输入Ⅸ因子的频率，但是目前这种基因治疗还不能达到满意的效果。

第三节 酶蛋白病

20世纪初英国医师Archibald Garrod最早研究了尿黑酸尿症、白化症、胱氨酸尿症和戊糖尿症。至今已发现数千种由于遗传性酶缺乏所引起的先天性代谢病。

一、氨基酸代谢病

氨基酸代谢病是指由于氨基酸分解代谢过程中酶的先天性缺乏而致氨基酸代谢缺陷。

(一) 苯丙酮尿症

苯丙氨酸(phenylalanine)是人体必需的氨基酸,它被用于制造黑色素、甲状腺素和肾上腺素等。代谢的每一步酶缺乏均可引起相关的疾病。

1934年Fölling首先发现苯丙酮尿症(phenylaketonuria,PKU)。群体发病率在欧美约为1/16 000。杂合子携带者约为1/70~1/50。患者家系中近亲婚配发病率可高达5%~14%。在智能低下的儿童中约0.5%~1.0%患此病。上海人群新生儿筛查阳性率为2/31 000,发生率与欧美接近。天津人群新生儿筛查阳性率为3/16 690,提示天津地区为本症高发区。

经典型苯丙酮尿症是由于肝脏中苯丙氨酸羟化酶(phenylalanine hydroxylase)缺乏使苯丙氨酸不能转变成酪氨酸,致使苯丙氨酸在体内累积,血清中苯丙氨酸浓度增高,患者血清中苯丙氨酸浓度可高达50~100mg/100ml(正常为1~3mg/100ml)。过量的苯丙氨酸使旁路代谢活跃,产生苯丙酮酸、苯乳酸、苯乙酸等(图8-12)。这些旁路代谢产物由尿液和汗液排出,使患儿的头发、皮肤和尿均有特殊的气味。过量的苯丙氨酸抑制酪氨酸脱羧酶的活性,影响去甲肾上腺素和肾上腺素的合成,也减少了黑色素的合成,使患者的毛发和肤色较浅。患者体内大量的苯丙氨酸竞争性地抑制色氨酸的羟化作用,同时它旁路代谢的产物抑制了5-羟色胺脱羧酶的活性,因此影响了色氨酸的正常代谢。旁路代谢产物堆积还抑制L-谷氨酸脱羧酶的活性,使γ-氨基丁酸生成减少,5-羟色胺(5-hydroxytryptamine,5-HT)和γ-氨基丁酸(γ-amino-butyric acid)减少,导致脑发育障碍。

患儿出生时无显著异常,3~4个月时逐渐出现症状,智力发育障碍,未予治疗者85%将发展到白痴水平。半数左右患儿有锥体外系损害症状,易激动(irritability),好动(hyperactivity),肌张力高(hypermyotonia),共济失调(ataxia),震颤(tremor)。约25%有惊厥(convulsion),多数有脑电图异常,骨骼发育迟缓,门齿稀疏。患儿可有呕吐(vomiting),且较严重,可误诊为幽门狭窄。皮肤、毛发和眼睛颜色变浅。小便有特殊的臭味(霉臭或鼠臭)。

本症为常染色体隐性遗传。苯丙氨酸羟化酶基因定位于12q24.1,长约90kb,有13个外显子。基因突变涉及核苷酸取代和缺失,绝大部分苯丙氨酸羟化酶分子缺陷来自核苷酸取代和缺失(表8-6)。

表8-6 苯丙氨酸羟化酶突变类型

苯丙氨酸羟化酶分子缺陷	患者人群
1. 核苷酸取代	
(1)甲硫1缬*(A→G)	法裔加拿大人
(2)精111终止(C→T)	中国人,未见于白人
(3)精158谷(G→A)	白人
(4)酪204半胱(A→G)	东方人
(5)精243终止(C→T)	白人
(6)精243谷胺(G→A)	东方人
(7)精252色(C→T)	白人
(8)精261谷胺(G→A)	白人
(9)谷280赖(G→A)	白人

续表

苯丙氨酸羟化酶分子缺陷	患者人群
(10)脯 81 亮(C→T)	白人
(11)亮 311 脯(T→C)	德国白人
(12)色 326 终止(G→A)	东方人
(13)酪 356 终止(T/C→A/G)	东方人
(14)精 408 色(C→T)	白人
(15)精 413 脯(G→C)	东方人
(16)酪 414 半胱(A→G)	白人
(17)3'IVS-4 拼接受者部位(AG→AA)	东方人
(18)5'IVS-12 拼接供者部位(GT→AT)	白人患者中最多见
2. 缺失	
(1) 缺失外显子 3	也门犹太人

* 表示第 1 位甲硫氨酸突变成缬氨酸，余类同

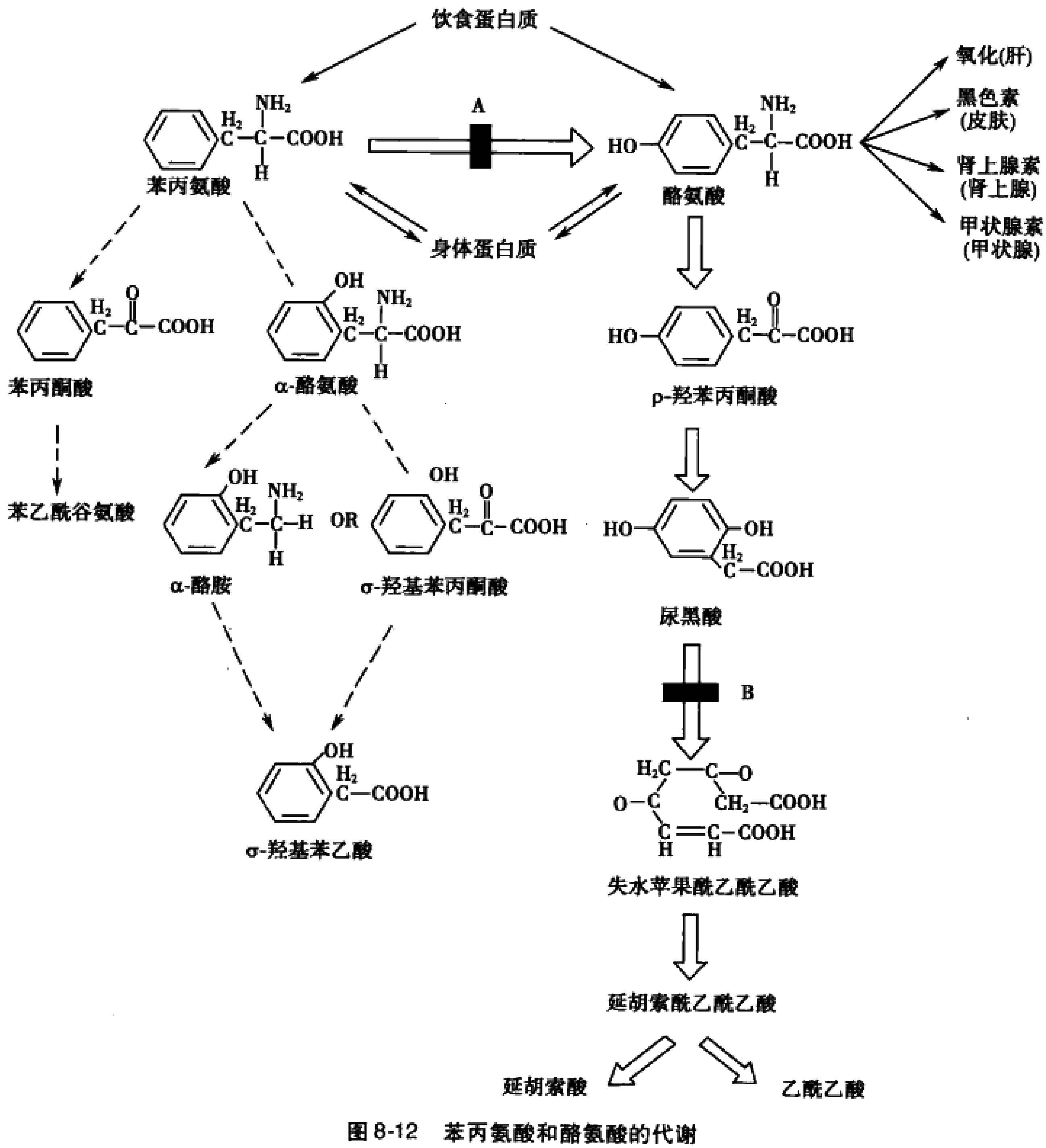


图 8-12 苯丙氨酸和酪氨酸的代谢

A：苯丙氨酸羟化酶缺乏导致苯丙酮尿症；B：是由于尿黑酸氧化酶缺乏导致尿黑酸尿症

苯丙氨酸羟化成酪氨酸需要一个辅助因子叫四氢生物蝶呤(tetrahydrobiopterin,XH_4)。它在苯丙氨酸羟化反应中,转变为醌式二氢生物蝶呤(quinoid dihydro-biopterin,XH_2),然后再在二氢蝶啶还原酶(dihydropteridine reductase)的催化下还原为四氢生物蝶呤(图8-13)。二氢蝶啶还原酶缺乏将使四氢生物蝶呤缺乏,从而虽有苯丙氨酸羟化酶存在也不能将苯丙氨酸羟化为酪氨酸,致使苯丙氨酸在体内积累,引起严重的苯丙酮尿症,被称为恶性或非典型性苯丙酮尿症。四氢生物蝶呤也是酪氨酸羟化成多巴(3,4-dihydro-phenylalanine,DOPA)最后形成黑色素和肾上腺素以及色氨酸羟化成5-羟色胺所必需的辅助因子。所以二氢蝶啶还原酶缺乏或二氢叶酸还原酶缺乏除了使血中苯丙氨酸增高外,同时使多巴、多巴胺、5-羟色胺、儿茶酚氨等浓度降低,从而引起一系列神经系统症状。已知二氢蝶啶还原酶基因定位于4p15.31。

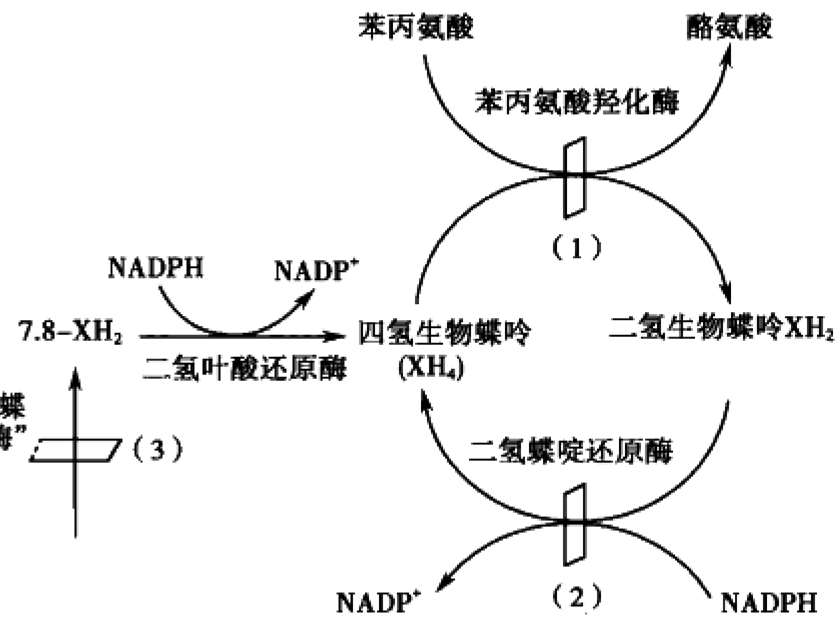


图8-13 苯丙氨酸羟化反应系统及其辅助因子四氢生物蝶呤的生成

(1)苯丙氨酸羟化酶缺乏引起经典型苯丙酮尿症;(2)二氢碟啶还原酶缺乏引起恶性苯丙酮尿症;(3)可能存在生物蝶呤合成酶缺乏

恶性苯丙酮尿症的临床表现与经典型苯丙酮尿症十分相似,多数也在出生后4~7个月时出现症状,表现有智能障碍,毛发肤色浅淡等。此外,还可表现有出生时体重低,头围小,流涎及不明原因的高热。肌张力低几乎是本病的特点。

经典型苯丙酮尿症患儿出生时无明显症状,通过实验室检查可明确诊断。新鲜尿液中加入$FeCl_3$可与尿中苯丙酮酸反应,形成绿色环。此方法简便,可用于筛选,但往往可能漏检。较为正确的方法是检查血中苯丙氨酸水平,患者血中苯丙氨酸含量增高。恶性苯丙酮尿症诊断可用高压液相色谱法(high pressure liquid chromato-graphy,HPLC)测定尿蝶呤谱,从中计算生物蝶呤占总蝶呤的百分值,还原酶缺乏时生物蝶呤百分含量较高,经典型苯丙酮尿症此百分值正常。

如能早期明确诊断,给予经典型苯丙酮尿症患儿以低苯丙氨酸饮食,可使智力发育正常。目前倾向于这种低苯丙氨酸饮食宜终生维持,因数据提示成年患者终止低苯丙氨酸饮食后仍可出现智力损害。恶性苯丙酮尿症患者即使给予低苯丙氨酸饮食也不能改善神经系统症状,故称之为恶性苯丙酮尿症。最近人工合成四氢生物蝶呤,用四氢生物蝶呤单一治疗或配合左旋多巴(levodopa)、5-羟色胺再加上脱羧抑制剂联合治疗使恶性苯丙酮尿症也不再是恶性的了。

1975年起,在美国43个州中,苯丙酮尿症的普查已法律化,95%以上的新生儿接受了检查,三年后,接受普查的新生儿超过1300万,发现1100名苯丙酮尿症患儿,并对他们进行了及时的治疗。我们国家在一些大城市里也陆续开始在新生儿中进行苯丙酮尿症的筛查。

苯丙酮尿症的女性患者在未经治疗的情况下怀孕,母亲血液中高浓度的苯丙氨酸会对胎儿造成不利的影响,产生体重低下、小头畸形、智力发育迟滞等症状。用饮食治疗可能可以改善这一情况,但治疗必须从受孕前开始,并在母亲妊娠过程中始终控制血液苯丙氨酸的浓度才可能获得一个健康的婴儿。

（二）尿黑酸尿症

尿黑酸尿症（alcaptonuria）是由于尿黑酸氧化酶（homogentisic acid oxidase）先天性缺乏所致，此酶主要存在于肝脏和肾脏。由于尿黑酸氧化酶缺乏，尿黑酸不能被最终氧化成乙酰乙酸和延胡索酸，结果大量尿黑酸从尿中排出。

新生儿期，生后不久发现尿布中有紫褐色斑点，洗不掉，日久渐使尿布呈黑褐色。

儿童期，尿黑酸尿是唯一的特点，此期除尿中排出尿黑酸之外，并无其他症状。

成人期，主要表现为尿黑酸尿、褐黄病和褐黄性关节炎。机体中尿黑酸增多，在结缔组织沉着，导致褐黄病（ochronosis）。多在 20 岁以后发病。在皮肤、耳廓、面颊、巩膜等处可见弥漫性色素沉着。此症如累及关节，则进展为关节炎，称褐黄病性关节炎（ochronoticarthritis）。

本症为常染色体隐性遗传，发病率大约为 1/250 000，携带者可用尿黑酸负荷试验检出。已知尿黑酸氧化酶基因定位于 3q21-q23。

大剂量使用维生素 C（抗坏血酸）可以降低结缔组织和尿黑酸的结合，饮食疗法减少酪氨酸和苯丙氨酸的摄入可以间接减少尿黑酸的排泄。

（三）眼皮肤白化症

眼皮肤白化症（oculocutaneous albinism）有Ⅰ型和Ⅱ型两种亚型，Ⅰ型为酪氨酸酶阴性，Ⅱ型为酪氨酸酶阳性。Ⅰ型为常见型，系由于遗传性酪氨酸酶（tyrosinase）缺乏，酪氨酸酶的作用是把多巴转化为黑色素。Ⅱ型的生化病理未明，患者可合成褐黑色素（pheomelanin），但真黑色素（eumelanin）的生物合成被阻断。

眼皮肤白化症为全身性白化。患者全身皮肤、头发、眼缺乏黑色素，故皮肤白皙，头发呈淡黄色，眼呈浅蓝色、畏光（photophobia）、视物模糊、可有眼球震颤（nystagmus）。日晒皮肤易灼伤，暴露的皮肤易患皮肤癌。此症的发病率为 1/35 000 ~ 1/15 000。

Ⅰ型和Ⅱ型在临床上难于区分，色素痣是鉴别Ⅱ型酪氨酸酶阳性眼皮肤白化症的唯一线索。

本症为常染色体隐性遗传。酪氨酸酶基因定位于 11q14-q21。Ⅰ型和Ⅱ型不是等位关系，两者受控于不同位点，可以互补。Trevor-Roper（1952）就是在发现了一对夫妻均为全身白化者却有四个肤色正常的孩子后，才揭示了有酪氨酸酶阳性眼皮肤白化症的存在。

二、糖代谢病

（一）半乳糖血症

半乳糖代谢涉及半乳糖激酶（galactokinase）、半乳糖-1-磷酸尿苷转移酶（galactose-1-phosphate uridyltransferase）和尿苷二磷酸半乳糖-4-表异构酶（uridine diphosphate galatose-4-epimerase）（图 8-14），这三种酶均检出有遗传性缺乏，导致不同亚型的半乳糖血症（galactosemia），均为常染色体隐性遗传。

半乳糖血症Ⅰ型为经典半乳糖血症，系由于半乳糖-1-磷酸尿苷转移酶遗传性缺乏所致。由于半乳糖-1-磷酸尿苷转移酶缺乏，致使半乳糖-1-磷酸在脑、肝、肾等处累积，导致损伤而致病。白内障的产生是由于半乳糖累积在晶体内，在醛糖还原酶的作用下转变成半乳糖醇（galactilol），后者改变晶体渗透压，使水分渗入晶体，致使晶状体变性。

患儿出生后喂乳汁（母乳、牛奶、羊奶等）后几天即出现呕吐、拒食、倦怠、腹泻、失重，1 周之后可表现出肝脏损害症状和黄疸、肝大、腹水（图 8-15）。1 ~2 个月内可出现白内障（cataract）。如不控制乳汁摄入，几个月后患儿出现智力发育障碍，肝脏损害更为严重，使凝血酶原缺乏而出血，低蛋白血症使全身水肿。还有生长发育障碍、蛋白尿和氨基酸尿。血和尿中半乳糖含量增高，而血糖低下。最终因肝功能衰竭或感染致死。

症状较轻者可以存活，而在出生 3 周后即可出现白内障，在 6 个月内可因门脉高压而致肝脾肿大。少数患者智力正常，但有肝硬化和白内障。

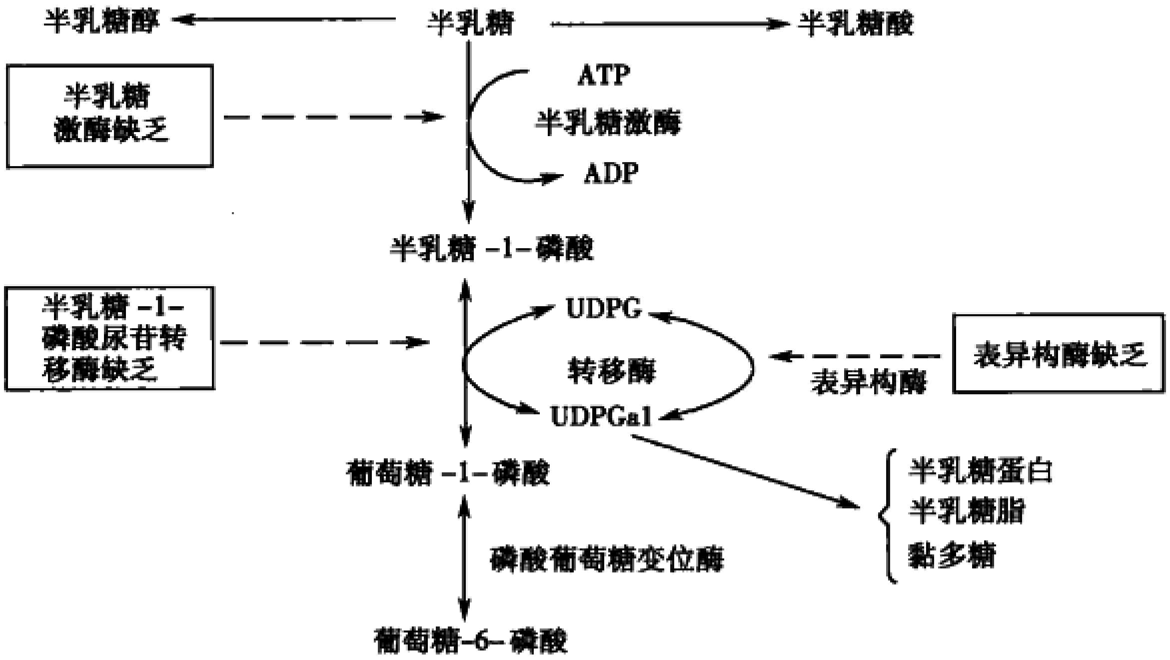


图 8-14 半乳糖代谢途径

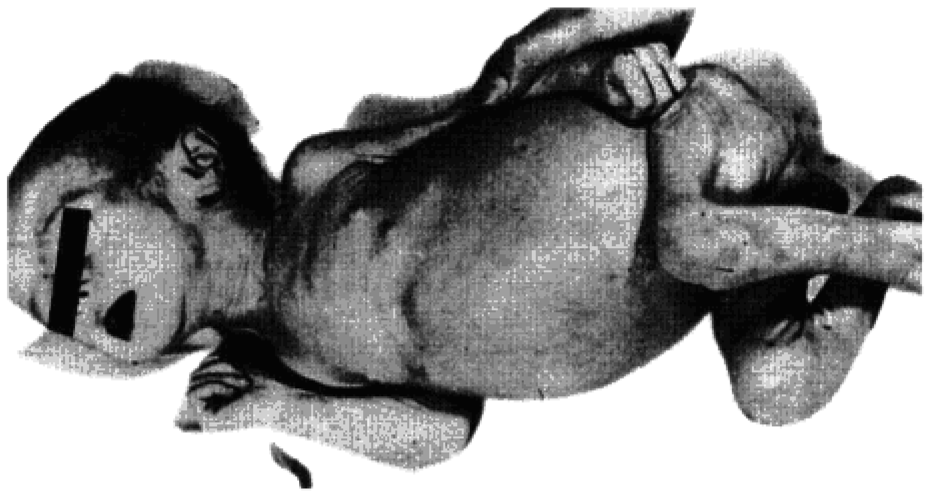


图 8-15 未经治疗的半乳糖血症婴儿示消瘦和腹水

婴儿出生后第一周末或第二周出现黄疸，伴有拒食或呕吐和肝脾肿大，应即怀疑患有本症。

半乳糖-1-磷酸尿苷转移酶的基因定位于9p13，此基因位点具有多态性，各等位基因的酶活性和电泳迁移率不同（表8-7）。多种基因型的杂合子显示有明显的剂量效应（表8-8）。一般杂合子只有正常纯合子酶活性的一半，而不呈现症状。

表 8-7 半乳糖-1-磷酸尿苷转移酶不同等位基因酶活性及电泳迁移率

等位基因	酶活性*(%)	电泳迁移率
正常 G^+	100	正常
半乳糖血症 G^G	0	—
Duarte G^D	50	快
Rennes G^R	10	慢
Indiana G^I	35	—，热不稳定
Los Angels G^L	100	快

* 每克血红蛋白每小时消耗 25μmol 尿苷二磷酸葡萄糖

半乳糖血症Ⅱ型为半乳糖激酶缺乏。除半乳糖尿和白内障与Ⅰ型相同外，本症有脑假瘤（pseudotumor cerebri），无氨基酸尿和蛋白尿，黄疸、肝肿大和智力发育障碍不常有。半乳糖激酶基因定位于17q24。

表8-8 半乳糖-1-磷酸尿苷转移酶的基因剂量效应

基因型	酶活性(%)	临床表现
G^+G^+	100	正常
G^+G^D	75	正常
G^DG^D	50	正常
G^DG^G	25	正常
G^IG^I	(35)	中度半乳糖血症
G^RG^R	10	半乳糖血症
G^GG^G	0	严重半乳糖血症

半乳糖血症Ⅲ型为尿苷二磷酸半乳糖-4-表异构酶缺乏，临床表现多变，可无临床症状或类似经典半乳糖血症。尿苷二磷酸半乳糖-4-表异构酶基因定位于1p36-p35。半乳糖血症三种亚型的临床症状比较见表8-9。

表8-9 半乳糖血症三种亚型的临床症状比较

半乳糖血症Ⅰ型(半乳糖-1-磷酸尿苷转移酶缺乏)	半乳糖血症Ⅱ型(半乳糖激酶缺乏)	半乳糖血症Ⅲ型(尿苷二磷酸半乳糖-4-表异构酶缺乏)
半乳糖尿	半乳糖尿	无临床症状或类似经典半乳糖血症
白内障	白内障	
黄疸	黄疸不常有	
肝肿大	肝肿大不常有	
智力障碍	智力障碍不常有	
氨基酸尿	无氨基酸尿	
蛋白尿	无蛋白尿	
拒食和呕吐	脑假瘤	
倦怠		
偶有腹泻		
肌张力低		
生长障碍		

本症可通过新生儿筛查发现患者。若能及时采取措施，严格限制婴儿饮食中的半乳糖成分，则症状可得到较好的控制。本症发病率各地统计有较大差别，英美和加拿大的统计为1/6万～1/4万。杂合子频率G^+G^G为0.9%～1.25%，G^+G^D为8%～13%。此症在我国的发病率不高，上海市新生儿筛查3万例中没有检出患者。

（二）黏多糖累积症

黏多糖(mucopolysaccharide)是由蛋白质和氨基多糖构成的糖蛋白。氨基多糖属直链杂多糖，多数是由糖醛酸和氨基己糖组成二糖单位，再重复连接长链，因含有较多的糖醛酸和硫酸基团，所以黏多糖呈酸性。大多数的氨基多糖通过木糖与蛋白质肽链的丝氨酸残基相连接。几种不同的氨基多糖链可同时存在于一条蛋白质肽链上，还可进一步聚合成更大的分子，结构十分复杂。硫酸皮肤素(dermatan sulfate)和硫酸乙酰肝素(heparan sulfate)是结缔组织的成分，硫酸皮肤素主要分布于皮肤、韧带、动脉及心瓣膜，硫酸乙酰肝素主要分布于大动脉、肝、肺等。黏多糖累积症(mucopolysaccharidosis, MPS)是由于糖苷酶(glycosidase)或硫酸酯酶(sulfatase)遗传性缺乏，造成酸性黏多糖部分分解产物在各种组织中累积而致病。患者面容粗犷，骨骼畸形，有的可有智力障碍和肝、脾、心等多种器官损害。根据缺乏酶的种类不同，黏多糖累积症可分为七型，见表8-10。

表 8-10 黏多糖累积症分型

病名	临床表现	酶缺乏	尿中过量的 MPS	遗传方式
MP S Ⅰ-H(Hurler 综合征)	角膜混浊,侏儒,骨骼异常,关节僵硬,智能发育不全,10 岁前死亡	α-L-艾杜糖苷酸酶	硫酸皮肤素 硫酸乙酰肝素	AR
MP S Ⅰ-S(Scheie 综合征)	角膜混浊,可有关节僵硬,主动脉瓣病,智力正常,寿命正常(?)	α-L-艾杜糖苷酸酶	硫酸皮肤素 硫酸乙酰肝素	AR
MP S Ⅰ-H/Ⅰ-S(Hurler/Scheie 复合综合征)	介于Ⅰ-H 和Ⅰ-S 之间	α-L-艾杜糖苷酸酶	硫酸皮肤素 硫酸乙酰肝素	AR
MP S Ⅱ-A 型(Hunter 综合征重型) MP S Ⅱ-B 型(Hunter 综合征轻型)	无角膜混浊,症状较Ⅰ-H 轻,通常 15 岁前死亡 轻微角膜混浊,智力尚可,可活到 30~60 岁	硫酸艾杜糖醛酸硫酸酯酶	硫酸皮肤素 硫酸乙酰肝素	XR
MP S Ⅲ-B 型(Sanfilippo B 综合征)	躯体改变较轻,中枢神经受损严重	N-乙酰 α 氨基葡萄糖苷酶	硫酸乙酰肝素	AR
MP S Ⅳ 型(Morquio 综合征)	特殊类型严重骨骼变化,角膜混浊,主动脉回流	硫酸软骨素硫酸 N-乙酰己糖胺硫酸酯酶	硫酸角质素	AR
MP S Ⅴ 型	以前指 Scheie 综合征			
MP S Ⅵ 型(Maroteaux-Lamy 综合征)	严重的骨骼变化,角膜改变,心瓣膜病,白细胞有包涵体,智力正常。轻型症状轻微	芳香基硫酸酯酶 B	硫酸皮肤素	AR
MP S Ⅶ型(Sly 综合征)(β-葡萄糖苷酸酶缺乏症)	肝脾肿大,多发性骨发育不全,白细胞包涵体,智力落后	β-葡萄糖苷酸酶	硫酸皮肤素	AR

1. 黏多糖累积症Ⅰ-H 型(Hurler 综合征) 黏多糖累积症Ⅰ-H 型为黏多糖累积症中的常见类型。患者面容粗犷(图 8-16),眉毛浓而连眉或称一字眉(synophrys),眼距增宽,鼻梁平塌,鼻孔宽而前倾,张口,唇舌大,牙小而疏,皮肤粗糙。骨骼异常,胸部畸形,胸腰部驼背,四肢短,掌宽而手指粗短,手指部分屈曲,似爪状手,关节僵硬,活动受限。呈现侏儒、渐进性智力发育不全和体质损害,进行性肝脾肿大,角膜混浊(corneal clouding)。患者由于溶酶体内缺乏 α-L-艾杜糖苷酸酶(α-L-iduronidase)致使硫酸乙酰肝素和硫酸皮肤素降解受阻而累积,尿中可检出过量的硫酸乙酰肝素和硫酸皮肤素。本症为常染色体隐性遗传。患者多数在 10 岁前死亡。新生儿中发病率为 1/100 000,杂合子频率为 1/150。



图 8-16 黏多糖累积症Ⅰ型(Hurler 综合征)患儿示粗犷面容

人们尝试了多种方法来治疗这种疾病,包括同种异体骨髓移植、α-L-艾杜糖苷酸酶的替代疗法、脐带血移植以及慢病毒载体转导表达人 IUDA 的 cDNA 进入患者的成纤维细胞。都取得了一定的疗效。

接受同种异体骨髓移植的患儿,平均五年生存率达到

了64%,能够达到延缓疾病进程、延长患者寿命的作用。并且接受移植时年龄小于2周岁者智力发育情况明显优于大于2周岁的患儿。由于能够找到相匹配骨髓供体的患者较少,人们用匹配3个HLA标记的含有正常α-L-艾杜糖苷酸酶活性的脐带血移植来代替这种方法,也取得了不错的效果。

2. 黏多糖累积症Ⅰ-S型(Scheie综合征) Scheie综合征的症状类似Hurler综合征,都有角膜浑浊,骨骼异常较黏多糖累积症Ⅰ-H型轻,身材不矮小,智力一般可属正常范围。还可有多毛(hirsutism),口大。通常有肝大,而无脾肿大。可活到成年。本症为常染色体隐性遗传。患者缺乏α-L-艾杜糖苷酸酶,使硫酸乙酰肝素和硫酸皮肤素累积。

3. 黏多糖累积症Ⅰ-H/Ⅰ-S型(Hurler/Scheie复合综合征) Hurler和Scheie综合征的隐性基因为等位性质,故可有Ⅰ-H/Ⅰ-S型复合杂合子导致复合综合征,其症状介于Hurler综合征与Scheie综合征二者之间。α-L-艾杜糖苷酸酶基因定位于4p16.3。所有的黏多糖累积症均可通过对羊水细胞培养后作酶的活性测定来进行产前诊断。

4. 黏多糖累积症Ⅱ型(Hunter综合征) 黏多糖累积症Ⅱ型与Ⅰ-H型有相似之处,但Ⅱ型无角膜混浊或较轻微,可借此与Ⅰ型相鉴别,智力损害较轻。患者呈现关节僵硬,矮小,爪状手,骨骼变化虽与Ⅰ-H型类同,但Ⅱ型较轻,发病晚,进展较慢,肝脾肿大,多毛,面容粗糙,无驼背(图8-17)。Ⅱ-A型有智力落后。Ⅱ-B型无智力落后。本症为X连锁隐性遗传。患者缺乏硫酸艾杜糖醛酸硫酸酯酶,尿中可检出过量的硫酸皮肤素和硫酸乙酰肝素。硫酸艾杜糖醛酸硫酸酯酶基因定位于Xq27.3-q28。黏多糖累积症Ⅱ型可直接从羊水中作酶的活性测定来进行产前诊断。

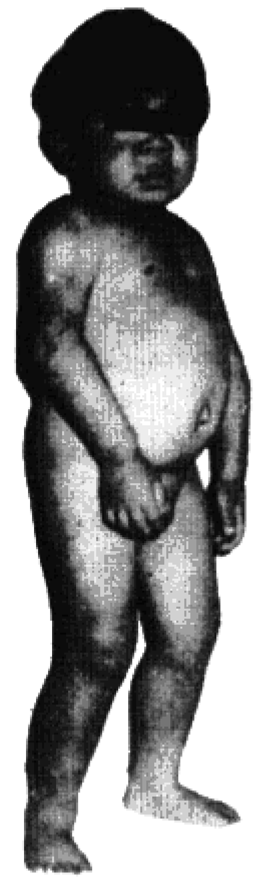


图8-17 黏多糖累积症Ⅱ型(Hunter综合征)患儿

三、脂类代谢病

脂类代谢病包括脂类分解代谢过程中特异性酶缺乏,导致其相应脂类底物在内脏、脑部和血管中累积,使这些系统混乱而致病,总称为脂类累积症(lipidosis)。脂类结构复杂,种类多样,特异性酶缺乏所致脂类累积症也多种多样,如神经鞘脂累积症(sphingolipidosis)、黏脂累积症(mucolipidosis)等。以下讨论限于神经鞘脂累积症的若干代表性病种。神经鞘脂(sphingolipid)的基本结构是酰基鞘氨醇(ceramide),后者系由鞘氨醇(sphingol, sphingosine)与脂肪酰长链相接而成。在鞘氨醇的第一位碳上可接上其他残基,构成神经鞘脂类化合物。如接上磷酸和胆碱,即为鞘髓磷脂(sphingomyelin),存在于脑和神经组织中。如接上葡萄糖,则为葡萄糖脑苷脂(glucocerebroside)。在接上多糖链后还可再接上1个或1个以上的N-乙酰神经氨酸(N-acetyl-neuraminic acid)〔一种唾液酸(sialic acid)〕,即为神经节苷脂(ganglioside)。脑苷脂和神经节苷脂均为糖脂(glycolipid),是神经髓鞘和膜结构的组成成分。正常情况下,神经鞘脂通过溶酶体中特异性水解酶水解而逐步分解,如特异性水解酶缺失,将使中间产物累积而致病(图8-18,表8-11)。

(一) Gaucher病

本症可分急性婴儿型和慢性型两种临床类型。

急性婴儿型主要有全身性肌张力过度,角弓反张(opisthotonos),斜视,吞咽困难,呕吐,喉头痉挛和呼吸困难,严重生长阻滞和恶病质,通常2岁前死亡。慢性型为隐匿性发作,最常见是肝脾肿大,贫血,最后可全血减少,多数患者有黄色或棕色斑块状色素沉着,体表暴露处尤为明显。成年患者常有结膜黄斑。急性婴儿型为葡萄糖脑苷脂酶(glucocerebrosidase)完全缺乏,使葡萄糖酰基鞘氨醇在网状内皮系统中累积,慢性型此酶有部分存在。在脾、肝、淋巴结、骨髓等处可见Gaucher细胞。这些细胞很大(直径20~100μm),有1个或1个以上小而致密的细胞核位于细胞周边。Gaucher细胞是本症诊断的病理学线索。本症为常染色体隐性遗传。葡萄糖脑苷脂酶是一种酸性β-葡萄糖苷酶,定位

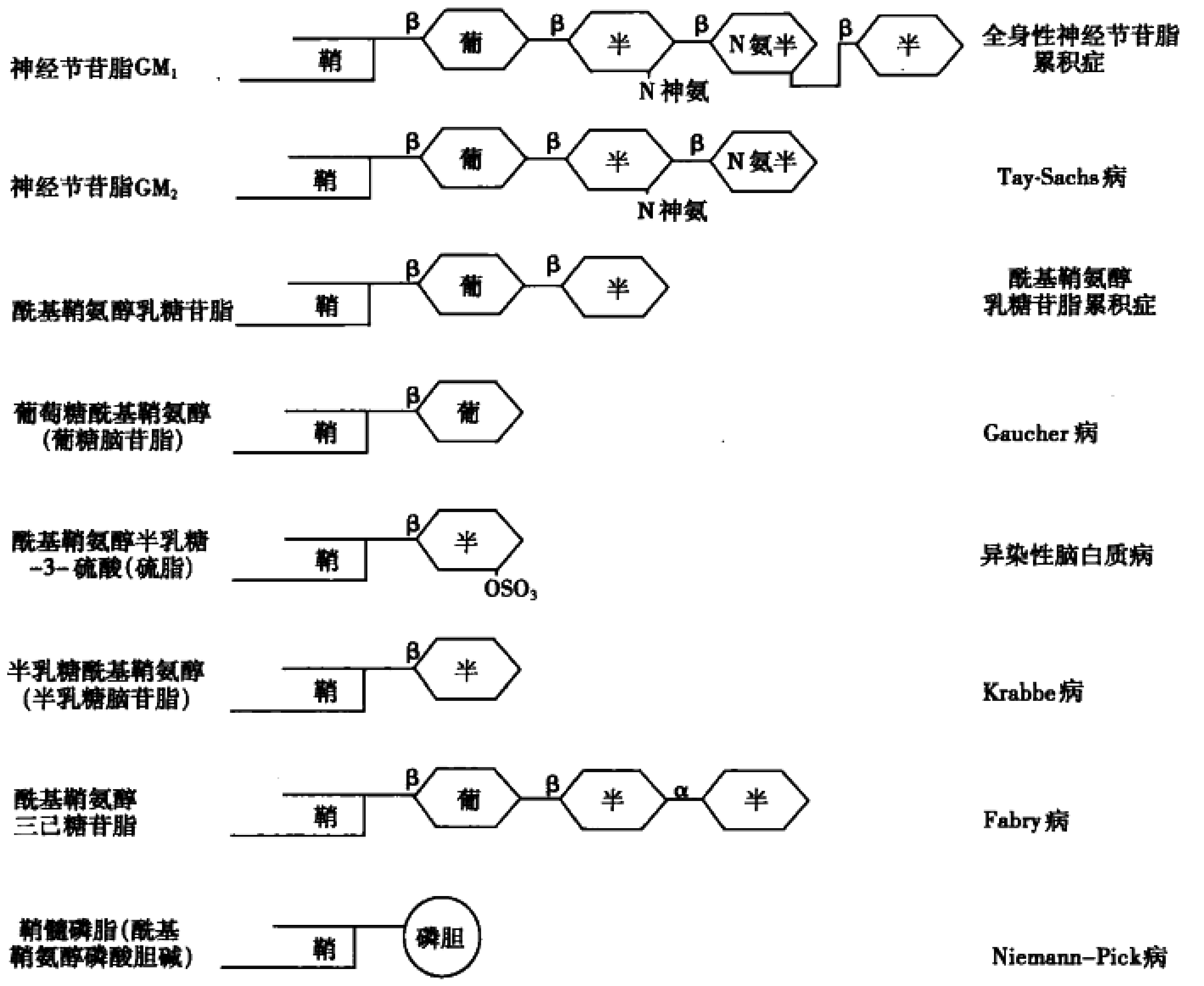


图 8-18 不同神经鞘脂累积症的主要神经鞘脂类的结构示意图

鞘:酰基鞘氨醇(ceramide) 葡:葡萄糖 半:半乳糖 磷胆:磷酸胆碱(phosphorylcholine) N神氨:N-乙酰神经氨酸(N-acetylneuraminic acid) N氨半:N-乙酰氨基半乳糖(N-acetylgalactosamine)

表 8-11 神经鞘脂累积症

病名	缺乏的酶	累积物	主要症状
全身性神经节苷脂累积症	β-半乳糖苷酶	GM_1	精神运动障碍,癫痫,失明。视网膜樱桃斑
Gaucher 病 Tay-Sachs 病(黑矇性白痴)	葡萄糖脑苷脂酶 氨基己糖苷酶 A	酰基鞘氨醇三己糖苷脂	肝脾肿大,骨髓等组织中有 Gaucher 细胞重度精神运动障碍,癫痫,失明。视网膜樱桃斑
酰基鞘氨醇乳糖苷脂累积症	酰基鞘氨醇乳糖苷脂 β-半乳糖苷酶	葡萄糖脑苷脂	临床症状各异
Sandhoff 病	氨基己糖苷酶 A 和 B	GM_2 Ⅱ型	同 Tay-Sachs 病,肝脾肿大,心肌病
Fabry 病	α-半乳糖苷酶		血管角化瘤,角膜混浊,肢端感觉异常,肾及心血管机能不全
Krabbe 病	半乳糖脑苷脂酶	半乳糖脑苷脂	重度神经运动障碍性痉挛,发音迟滞
异染性脑白质病	芳香基硫酸酯酶 A	酰基鞘氨醇半乳糖-3-硫脂硫酸	同上,进行性变性
Niemann-Pick 病	鞘髓磷脂酶	鞘髓磷脂	重度精神运动障碍,视网膜樱桃斑,肝脾肿大,骨髓有泡沫细胞等

于1q21。已检出的几种突变型见表8-12。采用等位基因特异性寡聚核苷酸(allele-specific oligonucleotide,ASO)探针检测370位G→A取代,并以RFLP检测444位T→C取代,可覆盖大约80%患者。可通过羊水细胞测定酶活性或检测分子变异进行产前诊断。

表8-12 Gaucher病患者葡萄糖脑苷脂酶(酸性β葡萄糖苷酶)变异体

精119谷胺*(G→A)	脯415精(C→G)
天胺370丝(A→G)	XOVR葡萄糖脑苷脂酶/假基因的融合基因(5'活性基因-3'假基因)
亮444脯(T→C)	

*表示第119位精氨酸突变成谷氨酰胺,余类同

(二) Tay-Sachs病

本症亦称GM_2神经节苷脂累积症(GM_2 gangliosidosis)和家族性黑朦性白痴(familial amaurotic idiocy)。氨基己糖苷酶A(hexosaminidase A)催化分解GM_2神经节苷脂成为GM_3和N-乙酰氨基半乳糖(N-acetylgalactosamine)(图8-19)。此酶缺乏时导致GM_2神经节苷脂累积。

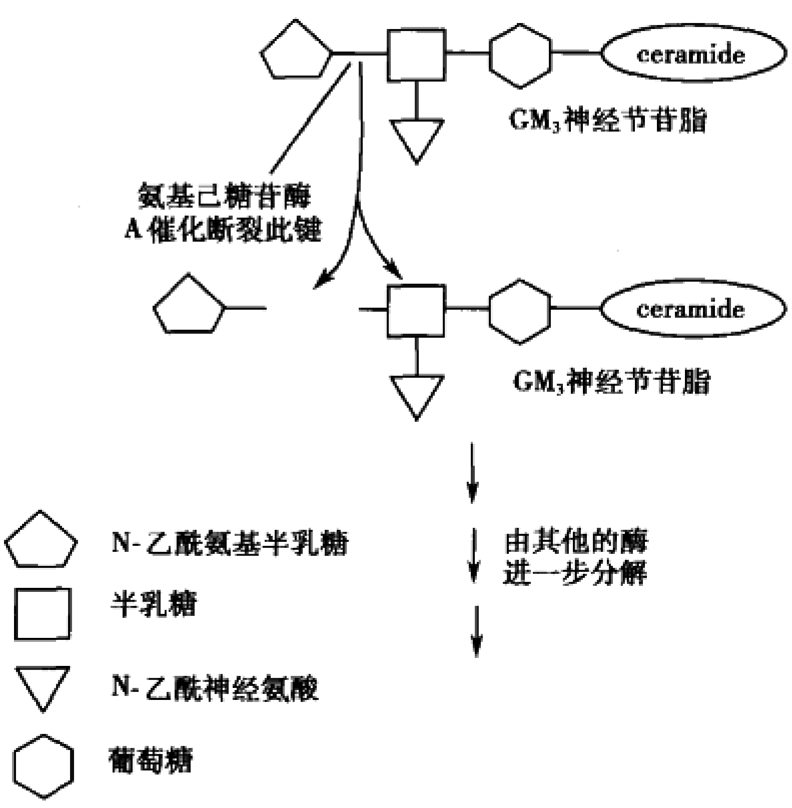


图8-19 氨基己糖苷酶A分解GM_2神经节苷脂
ceramide:酰基鞘氨醇

本症起病时最常见的症状为听觉过敏(hyperacusis)。病程早期即可见视网膜黄斑变性(macular degeneration),视网膜呈樱桃红斑点,进行性失明。可有抽搐,在局部性或全身性抽搐前数月可发生无动因的大笑。患儿进行性肌张力减退,衰弱,生长阻滞,到后期完全不能动弹,出现恶病质(cachexia),平均存活25.9个月。本症为常染色体隐性遗传,氨基己糖苷酶A基因定位于15q23-q24。已检出的氨基己糖苷酶A基因突变类型有核苷酸取代、缺失和移码。本症在阿什克奈兹族(Ashkenazi)犹太人中发病率最高。

(三) Niemann-Pick病

患者肝脾肿大,有全身性或局部性淡褐色黄斑,部分患者在口腔黏膜上有蓝色色素斑,斑点中有樱桃红小点。神经系统受损,有痉挛状态,最后成为白痴。发病较早,通常在出生1个月内发病。本症为常染色体隐性遗传。鞘髓磷脂酶(sphingomyelinase)基因定位于11p15.4-p15.1。可通过羊水细胞的酶活性测定来进行产前诊断。

四、嘌呤代谢病

在嘌呤代谢过程中已知有多种酶缺乏。完全缺乏次黄嘌呤鸟嘌呤磷酸核糖基转移酶(hypoxanthine guanine phosphoribosyl-transferase,HGPRT)导致Lesch-Nyhan综合征,此酶部分缺乏可致尿酸产生过量和痛风。腺嘌呤磷酸核糖基转移酶(adenine phospho-ribosyltransferase)缺乏可导致肾结石。黄嘌呤氧化酶(xanthine oxidase)缺乏导致黄嘌呤尿(xanthinuria),可有尿路黄嘌呤结石和肌痛(myalgia)。腺苷脱氨酶(adenosine deaminase,ADA)缺乏与严重联合免疫缺乏综合征相关联。肌腺苷酸脱氨酶(myoadenylate deaminase)缺乏的患者可表现虚弱,在剧烈运动后肌肉痉挛,静脉血氨不升高。

Lesch-Nyhan 综合征

本征亦称自残综合征(self-mutilation syndrome)。患者遗传性缺乏次黄嘌呤鸟嘌呤磷酸核糖基转移酶(hypoxanthine guanine phosphosibosyl-transferase,HGPRT)。此酶催化5-磷酸核糖-1-焦磷酸(5-phosphoribosyl-1-pyrophosphate PRPP)上的磷酸核糖基转移到鸟嘌呤和次黄嘌呤上,使之成为鸟嘌呤核苷酸(guanylic acid)和次黄嘌呤核苷酸(肌苷酸)(inosinic acid),而这两种核苷酸和腺嘌呤核苷酸可反馈抑制嘌呤前体5-磷酸核糖-1-胺(5-phosphoribosy-1-amine)的生成(图8-20)。如此酶缺乏,则鸟苷酸和肌苷酸合成减少,反馈抑制减弱,嘌呤合成加快,致使尿酸增高,代谢紊乱而致病。患者红细胞和白细胞中此酶含量可减少到正常人的2% ~10%。

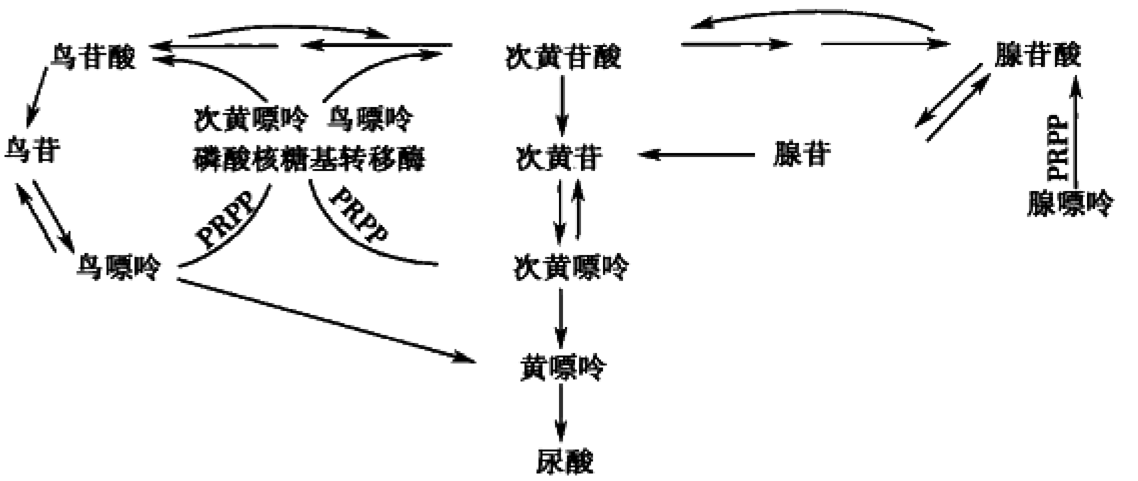


图8-20 嘌呤的生物合成和转换
PRPP:5-磷酸核糖-1-焦磷酸

患者临床症状有高尿酸血(hyperuricacidemia)和尿酸尿(uricaciduria),导致痛风性关节炎,大脑瘫痪,智力迟钝、舞蹈样动作,有强迫性自残行为(咬嘴唇和手指)(图8-21)。当酶呈部分缺乏时,患者出现痛风,可无其他严重症状。Lesch-Nyhan 综合征患者可活至20岁左右,多死于感染和肾功能衰竭。

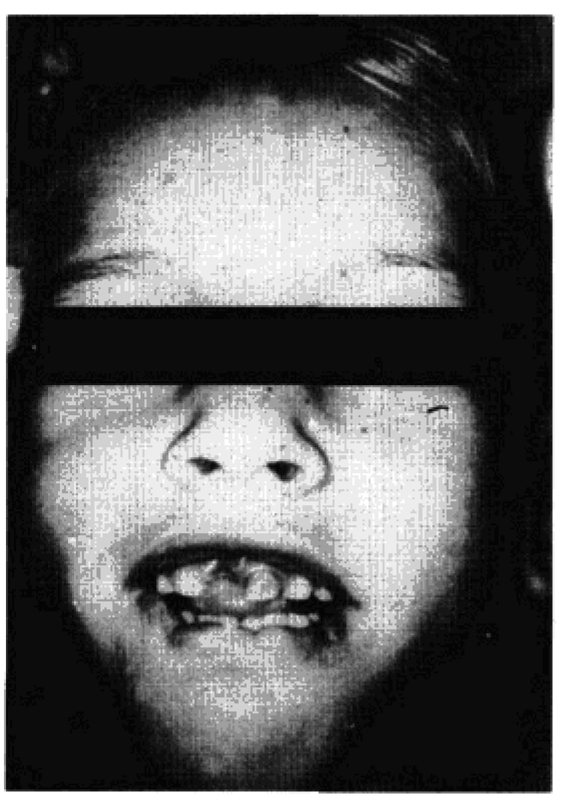


图8-21 Lesch-Nyhan 综合征
患儿示自残的下嘴唇

本征为X连锁隐性遗传。次黄嘌呤鸟嘌呤磷酸核糖基转移酶基因定位于Xq26-q27.2。已检出的此酶主要变异体的突变类型有核苷酸取代、插入、缺失和移码(表8-13)。可在DNA水平上作产前诊断。

五、卟啉代谢病

由于血红素合成过程中某种酶缺陷而引起的卟啉化合物或其前身物堆积,称为卟啉症。

血红素正常代谢过程中δ-氨基-γ-酮戊酸合成酶(δ-amino-laevu-linic acid synthetase,ALA合成酶)是血红素合成所需的第一个酶。正常情况下此酶的含量是较少的,它是整个代谢过程中控制速率的酶。血红素合成过程中δ-氨基-γ-酮戊酸以下任何一个反应中的酶缺乏均可使血红素合成减少,致使血红素对δ-氨基-γ-酮戊酸合成酶的负反馈受到抑制,从而使δ-氨基-γ-酮戊酸合成酶大量增加,中间代谢产物在体内累积,尿和粪便中有大量卟啉前体物排出,出现多种临床症状(图8-22)。

表 8-13 Lesch-Nyhan 综合征患者次黄嘌呤鸟嘌呤磷酸核糖基转移酶的突变类型

HGPRT 变异体	分子变异
1. 核苷酸取代	
(1)HGPRT(Detroit)	亮 41 脯*(T→C)
(2)HGPRT(New Haven)	甘 70 谷(G→A)
(3)HGPRT(Yale)	甘 71 精(G→C)
(4)HGPRT(Flint)	苯丙 74 亮(C→A)
(5)HGPRT(Arlington)	天冬 80 缬(A→T)
(6)HGPRT(Munich)	丝 104 精(C→T)
(7)HGPRT(London)	丝 110 亮(C→T)
(8)HGPRT(Midland)	缬 130 天冬(T→A)
(9)HGPRT(Ann Arbor)	异亮 132 甲硫(T→G)
(10)HGPRT(Milwauku)	丙 161 丝(G→T)
(11)HGPRT(Kinston)	天冬 194 天胺(G→A)
(12)HGPRT(New Briton)	苯丙 199 缬(T→G)
(13)HGPRT(Ashvine)	天冬 201 甘(A→G)
(14)HGPRT	内含子 7,5′供者部位(GTAAGT→GTAAAT)
(15)HGPRT	内含子 8,5′供者部位(G→A)
(16)HGPRT	内含子 1,3′受者部位(AG→TG)
(17)HGPRT	内含子 8,3′受者部位(AG→TG)
2. 插入	
(1)HGPRT(Evansville)	加长 24 个氨基酸
3. 缺失	
(1)HGPRT(Michigan)	缺失核苷酸 535～537(缬 179 缺)
(2)HGPRT(Connersville)	缺失核苷酸 532～609(苯丙 178～天胺 203)
4. 移码	
(1)HGPRT(Chicago)	核苷酸 56～58 间插入 1 个核苷酸(CCTTGA→CCTTTGA)

*表示第 41 位亮氨酸突变成脯氨酸,余类同

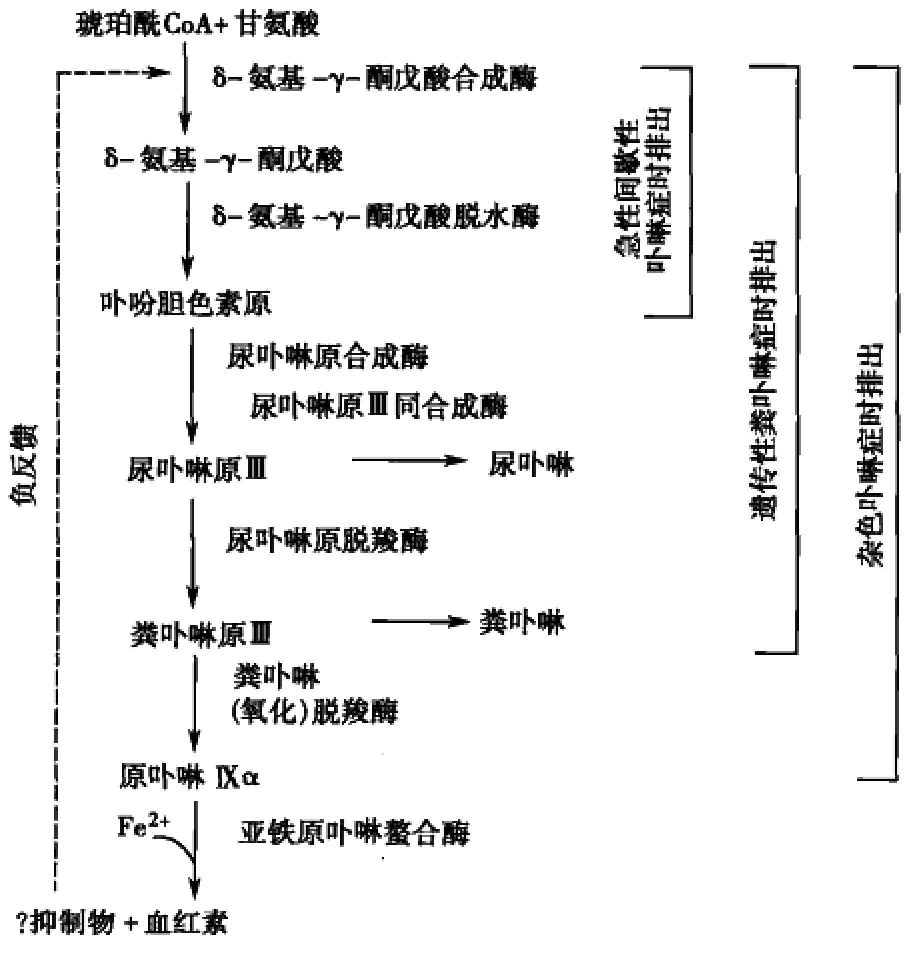


图 8-22 血红素的生物合成

由遗传性酶缺乏引起的卟啉症可分六型,急性间歇性卟啉症为其中之一。

急性间歇性卟啉症

急性间歇性卟啉症(acute intermittent porphyria)是肝卟啉症之一种。Strand 等(1970)认为急性间歇性卟啉症是由于肝细胞中尿卟啉原合成酶(uroporphyrinogen synthetase)的活性下降。尿卟啉原合成酶是在卟吩胆色素原(porphobilinogen)转化为尿卟啉原Ⅲ时的限速酶,但必须在尿卟啉原Ⅲ同合成酶(uroporphyrinogen Ⅲ cosynthetase)存在时才能合成尿卟啉原Ⅲ。患者由于尿卟啉原合成酶的活性下降,血红素合成减少,使负反馈受抑制,则 δ-氨基-γ-酮戊酸合成酶大量增加,使 δ-氨基-γ-酮戊酸及卟胆原的合成过量,大量在尿中排出(图 8-22)。Meyer 等(1972)证明患者虽不处在急性发作期,但红细胞中此酶的活性也同样降低,在发病时各种体细胞中还发现尿卟啉原合成酶活性降低。本症以腹绞痛(colic)、精神和神经症状间歇发作为特征。约 2/3 患者有肌肉麻痹(myoplegia),偶尔涉及锥体系的症状,有时可导致精神错乱。急性发作时或以腹部症状为主,或以神经症状为主。患者一生可发作一次或几次,持续几天到几个月。发病年龄在 18 ~ 20 岁左右。本症呈常染色体显性遗传。基因定位于 11q23.3。纯合子可能致死。患者为杂合子,但并非所有杂合子都有临床症状和生化异常。患者可为药物诱发,可能某些药物促发 δ-氨基-γ-酮戊酸合成酶活性增高。推测急性间歇性卟啉症发病系在尿卟啉原合成酶降低的基础上,再加上还原酶降低等其他代谢异常时,才可能在临床上出现明显的急性间歇性卟啉症。促使此症发病的一些常用药物见表 8-14。

表 8-14 促发遗传性肝卟啉症临床急性发作的药物举例

巴比妥类(barbiturates)	甲糖宁即甲苯磺丁脲(tolbutamide,orinase)
乙内酰脲(hydantoin)	麦角制剂(ergot preparations)
多睡酮即苯乙哌啶酮(glutethimide,doriden)	雌激素(estrogens)
灰黄霉素(griseofulvin)	孕激素(progestogens)
冬眠灵即氯丙嗪(chlorpromazine)	磺胺类(sulfonamides)

其他诱发因素还有酒精、感染、饥饿和激素水平的改变等等,并且在女性身上更易出现急性发作。产前诊断可以取成纤维细胞和羊水细胞经培养后筛查突变。

六、尿素循环代谢病

氨基酸的代谢产物,对人体属于无用的毒性物质,其中大部分经尿素循环变成尿素而排出(图 8-23)。尿素循环主要在肝脏中进行,涉及五种酶,其中任何一种酶的缺乏均影响尿素循环,使血氨升高,所有的临床症状均系血氨升高的结果。尿素循环中任何一步缺陷,均在新生儿期或幼婴期显示重症,有昏迷、抽搐或暴死。严重智力障碍是共同特征,尤其在吃了高蛋白食物后症状更为严重。

精氨酸血症

尿素循环中的三种氨基酸,鸟氨酸、瓜氨酸和精氨酸,仅起催化作用。氨基酸的量并不因参加尿素循环而增减。尿素循环中的最后一步是精氨酸在精氨酸酶的作用下使精氨酸分解,形成尿素(urea)和鸟氨酸(ornihine)。如精氨酸酶(arginase)缺乏,则血和脑脊液中精氨酸显著增高,而致精氨酸血症(argininemia)。

患者主要临床症状是血氨增高引起智力发育障碍、惊厥、嗜睡、呕吐。尿中有精氨酸、胱氨酸、赖氨酸、鸟氨酸和瓜氨酸排出。

本症为常染色体隐性遗传。精氨酸酶基因定位于 6q23。目前能用测定胎儿红细胞精氨酸酶活性的方法来进行产前诊断。

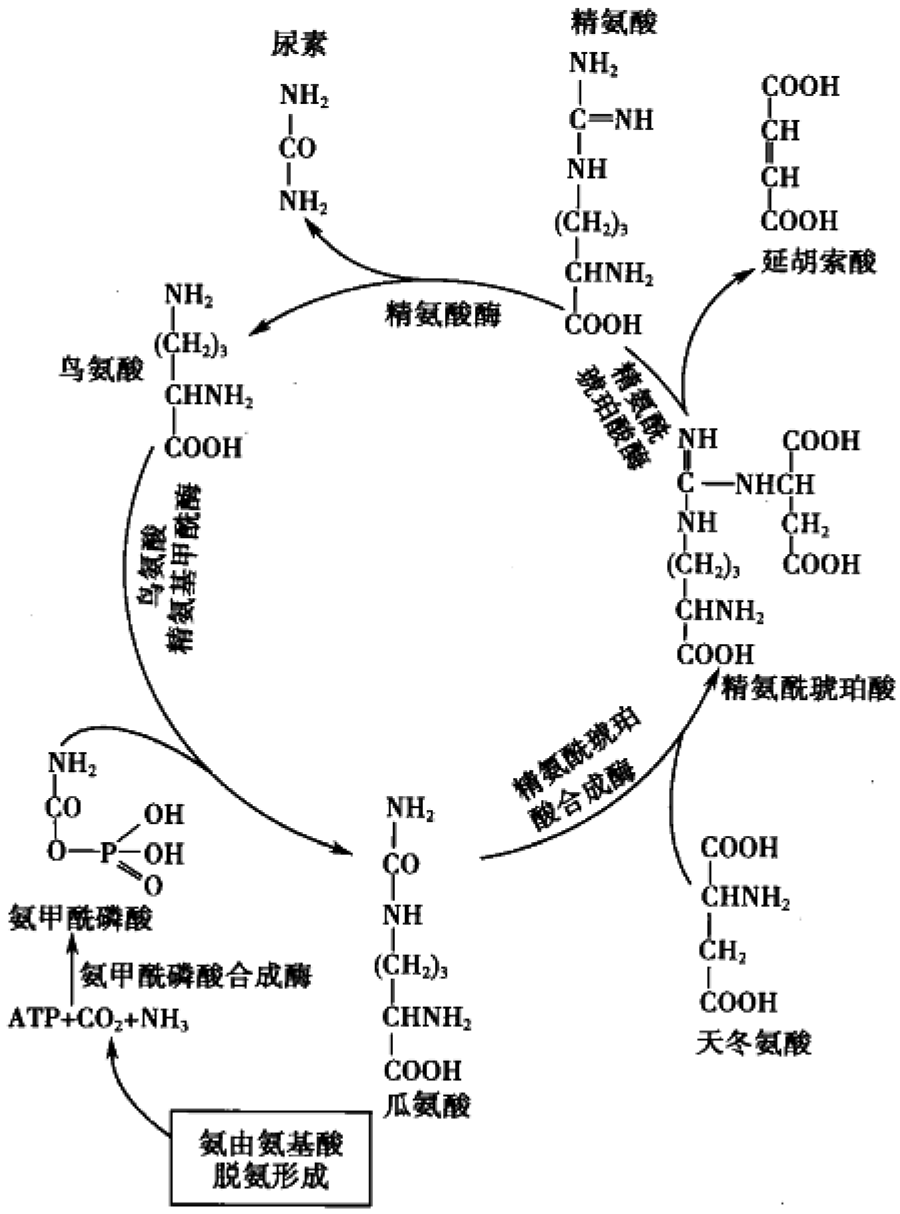


图 8-23 尿素循环中的酶

低蛋白饮食控制和苯甲酸钠的治疗可以减轻症状，减少血液中精氨酸的含量，智力发育可以高于同年龄其他患者。

第四节 受体蛋白病

细胞的代谢活动在相当程度上是由细胞的受体蛋白对外界分子的反应来调节。这些受体蛋白或是在细胞表面，或是在细胞质或细胞核内。有调节功能的信号分子都有它们特异的受体，包括多肽激素如胰岛素、高血糖素和促肾上腺皮质激素，固醇类激素如醛固酮、皮质醇和二氢睾酮，以及其他物质如一些血浆脂蛋白。这些信号分子结合到它们特异的受体上，会引起细胞一系列反应，特异地改变细胞的代谢格局，其中无疑涉及酶活性的改变。因此可以理解，一旦受体的生物合成有缺陷，合成了结构异常的受体分子，或是合成的数量大减甚至不能合成受体，就会对代谢过程带来复杂的干扰而致病。这些受体蛋白病比经典的单基因代谢更为复杂，研究自然也较为困难。Goldstein 等人自 70 年代以来长期从事家族性高胆固醇血症（familial hypercholesterolemia）低密度脂蛋白（low density lipoprotein，LDL）受体遗传的研究，取得丰硕成果。家族性高胆固醇血症无疑是研究其他遗传病类似受体缺陷的一个重要的模型。

家族性高胆固醇血症

遗传性高脂蛋白血症（hyperlipoproteinemia）患者血浆中胆固醇（cholesterol）和甘油三酯（triglyceride）增高，导致冠心病，心肌梗死。此症可有六种类型，分别决定于不同的位点，其中遗传性高脂蛋白

血症Ⅱ型为家族性高胆固醇血症，此症是由于细胞膜上低密度脂蛋白受体缺陷而致病。

在正常代谢中 LDL 同细胞膜上的 LDL 受体结合，通过内吞进入细胞，然后被溶酶体酸性水解酶水解，释放出游离胆固醇，后者在细胞内可激活脂酰辅酶 A：胆固醇脂酰转移酶（fatty acyl CoA：cholesterol acyl-transferase，ACAT）被酯化成胆固醇酯而贮存，同时游离胆固醇抑制 β-羟基-β-甲基戊二酰辅酶 A 还原酶（HMG CoA reductase）的活性，从而减少胆固醇的合成（图 8-24）。患者因 LDL 受体缺陷，LDL 不能进入细胞，使细胞内胆固醇的反馈抑制受阻，对 HMG CoA 还原酶抑制解除，致使细胞内胆固醇合成增高，血液及细胞内胆固醇堆积而致病。

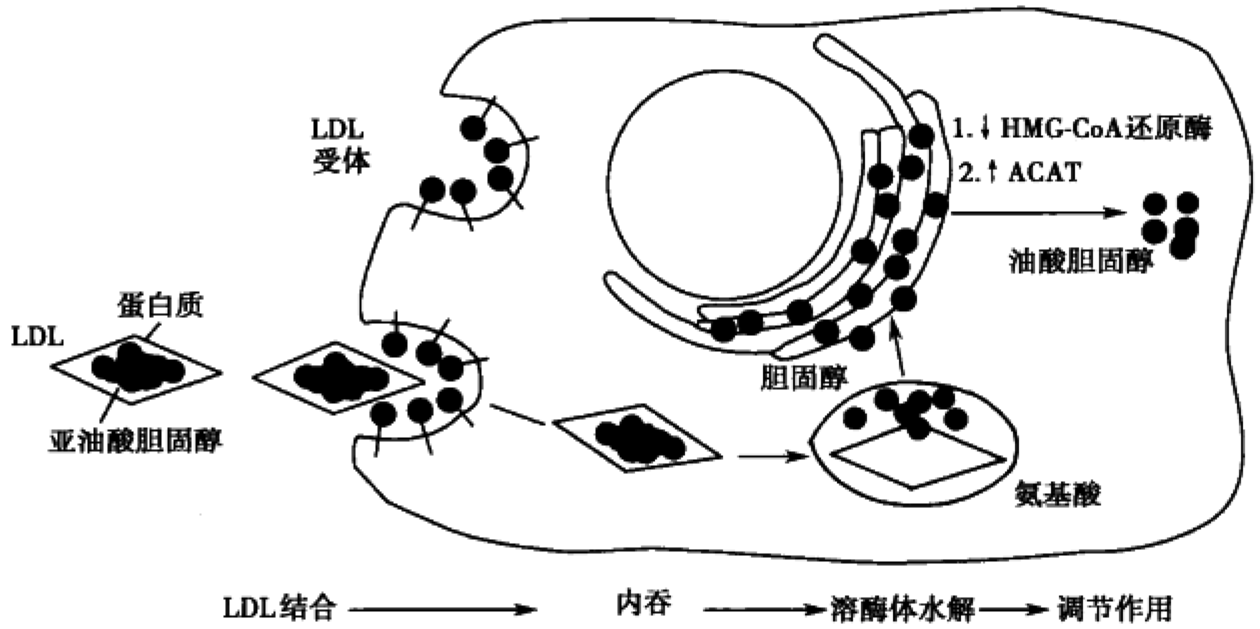


图 8-24　纤维母细胞低密度脂蛋白受体作用示意图
LDL：低密度脂蛋白　HMG-CoA 还原酶：β-羟基-β-甲基戊二酰辅酶 A 还原酶
ACAT：脂酰辅酶 A：胆固醇脂酰转移酶

本症为常染色体显性遗传。患者多为杂合子（图 8-25）。由于胆固醇沉积而出现黄瘤（xanthoma）随患者年龄增长而日益严重（图 8-26）。患者较早出现角膜弓（老人环）（arcus corneae），过早出现冠心病。纯合子患者病情更为严重，可在儿童期发生冠心病，5～30 岁即出现心绞痛（angina pectoris）和心肌梗死（myocardial infarction）的症状，可能骤死。杂合子发生率约为 1/500，产生正常人 40% 的 LDL 受体，血浆总胆固醇为 300～400mg/dl。纯合子发生率约为 1/1 000 000，只有正常人 10% 的 LDL 受体，血浆总胆固醇可高达 600～1200mg/dl。

一个成熟的 LDL 受体分子，氨基端是 LDL 结合区，羟基端有穿膜区和细胞质区。LDL 受体基因定位于 19p13.1-p13.2，有 18 个外显子。已检出多种变异体，有核苷酸取代、缺失、插入等，以缺失为多见。

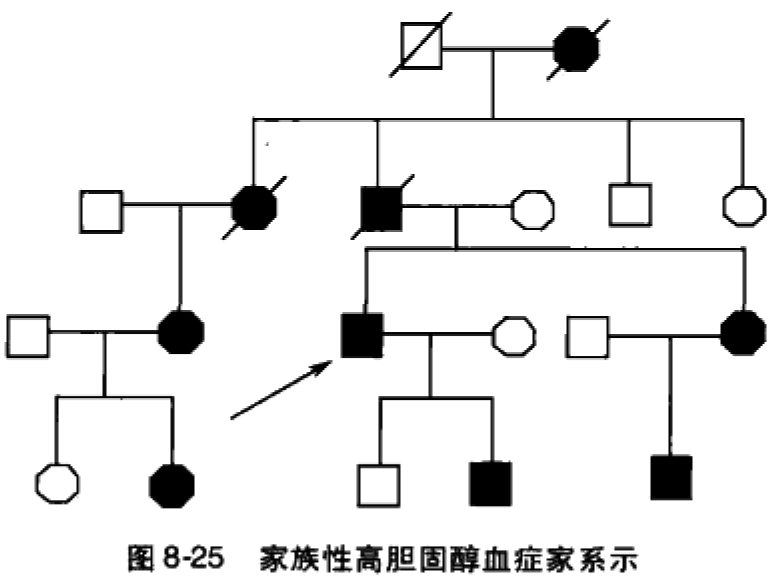


图 8-25　家族性高胆固醇血症家系示常染色体显性遗传

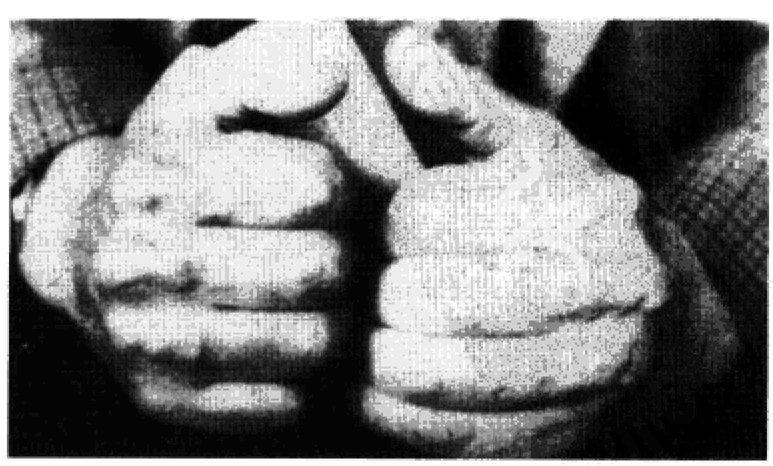


图 8-26　家族性高胆固醇血症患者手部示肌腱黄瘤

第五节　膜转运载体蛋白病

小分子量物质通过细胞膜而进出细胞，有相当程度取决于各种特异性的主动转运系统。例如小肠对各种小分子量物质的吸收往往只对一种物质或对少数有关物质是高度特异的，肾小管上皮对通过肾小球而流经肾小管的物质进行选择性重吸收、这些都属主动转运系统。在这种主动转运系统中如发生载体蛋白的遗传性缺陷，无疑也将致病。

一、肝豆状核变性

肝豆状核变性（hepatolenticular degeneration）是一种铜代谢障碍导致的疾病。本症由 Wilson 于 1912 年首次描述，故又称 Wilson 病。本症特点是铜过量沉积于组织中，尤其在肝、肾、脑、角膜等处。临床表现为进行性肝硬化（hepatocirrhosis），伴有基底神经节豆状核（lenticular nucleus）变性而引起的神经症状。

本症为常染色体隐性遗传，群体发病率约 30/1 000 000，杂合子频率约 1/90，致病基因频率为 0.56%，而撒丁岛发病率明显高于其他地区，每年都会有 10～12 个新发病例报道。本症发病的原因在于细胞膜与铜的转运有关的 ATP7B 的缺陷导致铜不能从细胞内及时清除。铜在组织中沉积所引起的毒性作用，在肝、肾、脑等组织特别明显。肝脏中铜的浓度可达正常的 5～30 倍，引起肝细胞坏死，发展为肝硬化。铜沉积于肾脏，使近曲小管受损而出现氨基酸尿（aminoaciduria）、蛋白尿（proteinuria）等，沉积于脑组织，引起神经系统的毒性反应；沉积于角膜，形成具有诊断意义的角膜外缘绿色环（Kayser-Fleischer 环）。铜被红细胞摄取后，可发生溶血性贫血。*ATP7B* 基因位于 13q14.3。

本症发病多在青少年期，大多在 6～20 岁之间。起病方式很不一致。40% 患者以肝损害为主，多发生在儿童，发病形式类似慢性活动性肝炎，后来发展到肝硬化，伴有蜘蛛痣（spider nevus，spider telangiecta-sis）、黄疸、肝脾肿大、腹水（ascites）、便血等肝功能不全的症状，肝功能试验常见阳性，如白：球比例倒置，絮浊反应异常，碱性磷酸酶中度增高，凝血酶原时间延长等；约有 40% 患者以神经症状为主，多发生在成年患者，有发音和吞咽困难、运动失调、体态异常、僵硬、震颤，偶有癫痫发作等；约有 20% 患者出现精神症状，有时可被误诊为精神分裂症，也有患者兼有神经损害和肝病表现。

头颅作计算机 X 线断层摄影（computerized tomography，CT）检查常有阳性发现，最常见者为脑萎缩，其余表现为豆状核低密度灶、脑干萎缩、脑室扩大、小脑萎缩、丘脑软化灶、尾状核低密度灶。肝型和脑型的 CT 改变无显著性差异。头颅 CT 常是诊断本症有价值的指标。

若在肝硬化或神经系统症状出现前就进行治疗，所有的症状均可得到控制。故人们用限制性片断长度多态性（RFLPs）进行症状前诊断，来改善这种疾病的预后，用 RFLP 间接标记通过连锁原理也可以对这种疾病作出产前诊断。

二、先天性葡萄糖、半乳糖吸收不良症

先天性葡萄糖、半乳糖吸收不良症（congenital glucose-galactose malabsorption）是由于患者小肠上皮对葡萄糖、半乳糖吸收过程中的载体蛋白有缺陷，所以对葡萄糖、半乳糖的吸收受阻。由于葡萄糖和半乳糖未被吸收，肠道内渗透压改变而使肠液增加，患者出现水样腹泻，与尿很相似，腹泻的发生和程度与喂给糖的时间及量有关，给食后 24 小时即出现腹泻，进食越多，越严重，但患儿进食量却又很大，所以体重很快下降。继而呈现脱水、失重、营养不良。患儿只要不以果糖作为主要糖类来源，就会出现腹泻，故要终身限制食用葡萄糖及半乳糖。但随着年龄增长对葡萄糖和半乳糖有所耐受。

目前认为本症属常染色体隐性遗传，相关基因定位于 22q13.1。

三、胱氨酸尿症

胱氨酸尿(cystinuria)是肾小管对四种氨基酸即胱氨酸(cystine)、赖氨酸(lysine)、精氨酸(arginine)和鸟氨酸(ornithine)的转运存在特异性的缺陷,使肾小管对肾小球滤出液中这四种氨基酸的重吸收出了问题,故患者血浆中这四种氨基酸水平低于正常值,而尿液中这四种氨基酸水平增高。

由于这四种氨基酸中赖氨酸、精氨酸和鸟氨酸均极易溶于水,而胱氨酸较不易溶于水,所以当患者每天排出量在0.5～1g胱氨酸时,尿中的浓度达到饱和,胱氨酸从尿液中析出,形成结晶。有人指出,胱氨酸尿症患者的小肠黏膜上皮细胞的主动转运也似有异常。

患者在生长或发育上通常都不会有任何营养性的异常。主要症状是尿路结石,可由此引起尿路感染和绞痛。

胱氨酸尿症至少涉及三个不同的异常等位基因,各以一种特异的方式影响主动转运过程,呈现三种不同的临床类型。

Ⅰ型纯合子患者四种氨基酸排泄量均增加,杂合子和正常纯合子四种氨基酸排泄量均正常。患者肠上皮对四种氨基酸均不吸收,为常染色体隐性遗传,相关基因定位于2p16.3。

Ⅱ型纯合子患者四种氨基酸排泄量均增加,杂合子的胱氨酸和赖氨酸排泄量有少量增加,正常纯合子排泄量正常。纯合子患者小肠上皮对四种氨基酸有少量摄取。此型呈现不完全隐性遗传,相关基因定位于2p16.3。

Ⅲ型与Ⅱ型相似,亦为不完全隐性遗传,其纯合子患者小肠上皮对氨基酸的吸收较为正常,但也低于正常人群。相关基因定位于19q13.1。

Summary

The hemoglubinnopathies were the first group of genetic diseases to be examined at a molecular level, because they exhibit such a diversity of molecular pathology. Most other genetic conditions have been found to result from mechanisms of altered gene function that have already been recognized, in some form, in a hemoglobin disorder. Thus although the hemoglobinopathies are paradigms for the genetic and mutational basis of "molecular disease", the biochemical pathology of each disease reflect different and often unique aspects of the structure and function of the protein altered by the mutation.

As the biochemical pathology of an increasing number of genetic diseases is gradually untangled, and as the genetic components of the common multifactorial diseases are characterized, new and unforeseen pathophysiological mechanisms will be recognized. The understanding of genetic disease at a molecular level not only contributes to the knowledge of normal human biology but is also the foundation of effective treatment for these disorders.

(陆振虞)

参考文献

1. Berg MA. Diverse growth hormone receptor gene mutations in Laron syndorme. Am J Hum Genet, 1993, 52: 998-1005
2. Eisensmith RC, Woo SL. Molecular basis of phenylketonuria and related hyperphenylalaninemias: mutations and polymorphism in the human phenylalanine hydroxylase gene. Hum Mutat, 1992, 1: 13-23
3. Green PM. Molecular pathology of hemophilia B. EMBOJ, 1989, 8: 1067-1072
4. Huang SZ, et al. The same TATA box β-thalassemia mutation in Chinese and US Blacks: Another example of independent or-

igins of mutation. Hum Genet, 1986, 74: 162-164

5. Valle DL, Mitchell GA. Inborn errors of metabolism in the molecular age. Prog Med Genet, 1988, 7: 100-129
6. Zhang YZ. Charaterization of genetic defects of hemophilia in mainland China. Genet Anal, 1999, 15(6): 205-207
7. Friedman JM. Genetics. 2ed edition, Williams & Wilkins, 1996. 115-117
8. Pasternak JJ. Human Molecular Genetics. Bethesda: Fitzgerald Science Press, 1999. 276-284
9. McKusick VA. Mendelian inheritance in man. 12th edition, Baltimore: The Johns Hopkins University Press, 1998
10. D. C. Ree. A Novel Silent Posttranslational Mechanism Converts Methionine to Aspartate in Hemoglobin Bristol(β67[E11] Val-Met→Asp). Blood, 1996, 88: 341-348
11. Hardison. HbVar: a relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. Hum Mutat, 2002, 19: 225-233
12. MannoCS. AAV-mediated factor IX gene transfer to skeletal muscle patients with severe hemophilia B. Blood, 2003, 101: 2963-2972
13. Mannucci PM. Treatment of von Willebrand's disease. New Eng J Med, 2004, 351: 683-694
14. Drogari E. Timing of strict diet in relation to fetal damage in maternal phenylketonuria. Lancet, 1987, 2: 927-930
15. Frances J R. New England Maternal PKU Project: Prospective study of untreated and treated pregnancies and their outcomes. J Pediat, 1987: 391-398
16. Charles Peters. Hurler Syndrome: Ⅱ. Outcome of HLA-Genotypically Identical Sibling and HLA-Haploidentical Related Donor Bone Marrow Transplantation in Fifty-Four Children. Blood, 1998, 91(7): 2601-2608
17. Kakkis ED. Enzyme-replacement therapy in Mucopolysaccharidosis I. New Eng J Med, 2001, 344(3): 182-187
18. Cossu P. Prenatal diagnosis of Wilson's diease by analysis of DNA polymorphism. New Eng J Med, 1992, 327: 357
19. Ala A. Wilson disease. Lancet, 2007, 369(9559): 397-408
20. Roberts EA. Liver as a key organ in the supply, storage, and excretion of copper. Am J Clin Nutr, 2008, 88(3) 851-854

第九章 线粒体遗传病

线粒体(mitochondria)是真核细胞的能量代谢中心。细胞呼吸作用中的氧化还原反应在线粒体中进行,并在此过程中产生大量三磷酸腺苷(ATP)作为能量供给整个机体利用。因此,线粒体被称为细胞的氧化中心和动力工厂。自 1894 年在动物细胞质内发现线粒体以来,人们对线粒体的结构、功能及其与疾病关系的认识不断深入。1963 年 Nass 首次在鸡卵母细胞中发现线粒体中存在 DNA;同年,Schatz 分离到完整的线粒体 DNA(mitochondrial DNA,mtDNA),从而开始了对 mtDNA 的探索。有性生殖中受精方式决定了线粒体遗传属母系遗传。早期曾有学者提出某些疾病可能是细胞质遗传所致,但直到 1987 年 Wallace 等通过研究 mtDNA 突变和 Leber 遗传性视神经病之间关系后,才明确地提出 mtDNA 突变可引起人类的疾病。近年来这一领域研究迅猛发展,目前已证实人类 100 多种疾病与 mtDNA 突变有关。

第一节 mtDNA 的结构特点与遗传特征

一、mtDNA 的结构特点

mtDNA 是指位于线粒体中的 DNA,而真核细胞中大部分 DNA 存在于细胞核中。线粒体是动物细胞核外唯一含有 DNA 的细胞器,它存在于除成熟红细胞外的所有组织细胞中。Anderson 等在 1981 年发表了完整的人类 mtDNA 序列。人类 mtDNA 是一个全长为 16 569 bp 的双链闭合环状分子(图 9-1),外环为重(H)链,内环为轻(L)链。人类 mtDNA 包含 37 种编码基因,分别编码 13 种多肽链、22 种 tRNA 和 2 种 rRNA。线粒体上的基因编码紧凑,没有内含子。mtDNA 唯一的非编码区是约 1000bp 的 D-环,它包含 mtDNA 重链复制起始点、轻重链转录的启动子以及 4 个高度保守的序列,分别位于 213 ~ 235、299 ~ 315、346 ~ 363 以及终止区 16 147 ~ 16 172。mtDNA 两条链的碱基组成差别较大,H 链富含 G,而 L 链多含 C。H 链是 12 种多肽链、2 种 rRNA 和 14 种 tRNA 转录的模板,而 L 链仅作为 1 种多肽链和 8 种 tRNA 转录的模板。mtDNA 具有两个复制起始点,分别起始复制 H 链和 L 链,其 H 链复制的起始点(OH)与 L 链复制起始点(OL)相隔 2 ~ 3 个 mtDNA。

与核 DNA 不同,mtDNA 分子上无核苷酸结合蛋白,缺少组蛋白的保护,基因与基因之间少有间

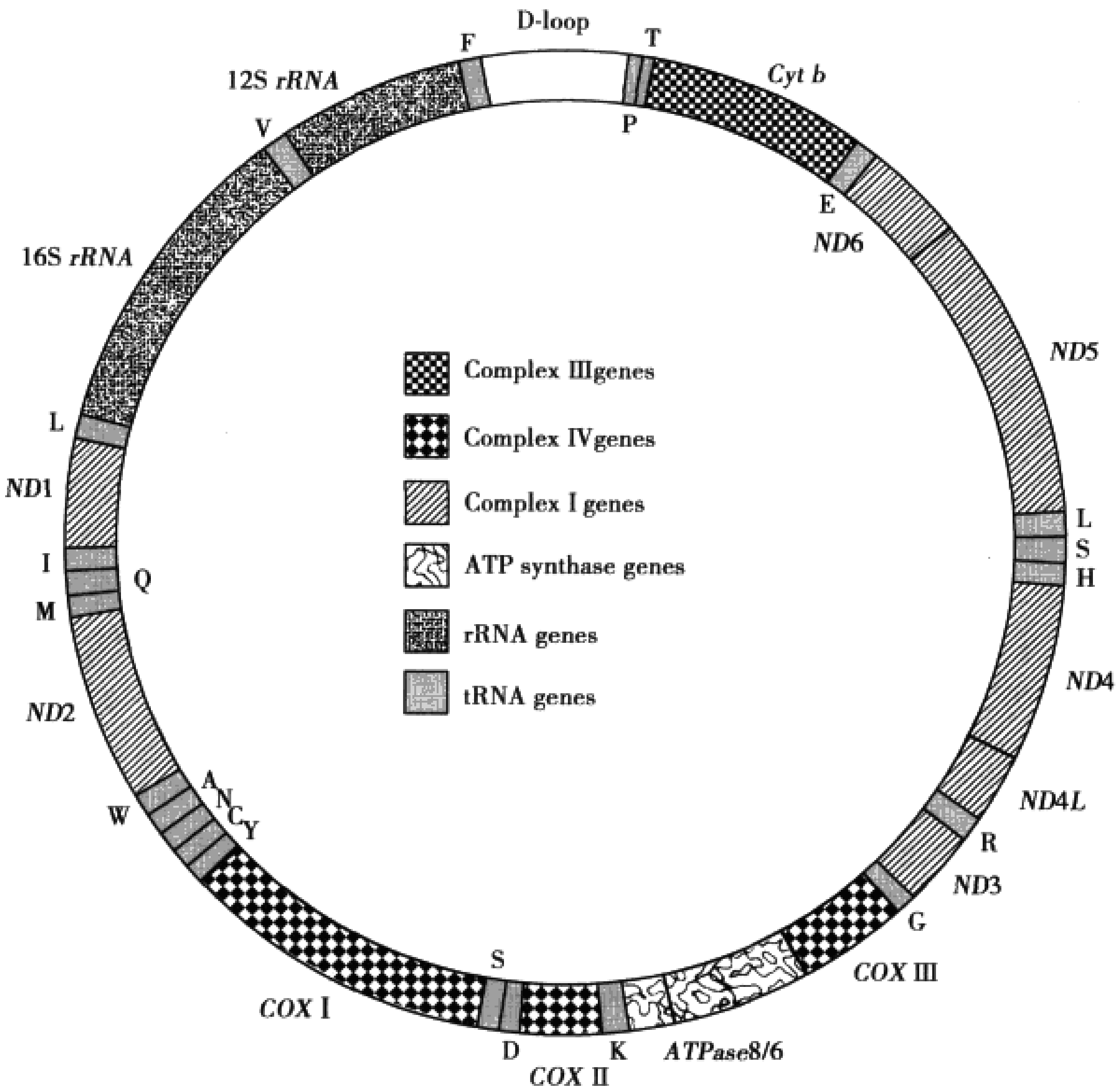


图 9-1 线粒体结构图

隔，而且线粒体内无 DNA 损伤修复系统，这些因素成为 mtDNA 易于突变且突变难以修复并遗传到子代细胞的分子基础。mtDNA 的另一特点是每一个细胞中含有数百个线粒体，每个线粒体内含有2～10 个拷贝的 mtDNA 分子，每个细胞可具有数千个 mtDNA 分子，从而构成细胞中 mtDNA 异质性的分子基础。

二、mtDNA 的遗传特征

与核 DNA 相比，mtDNA 具有其独特的传递规律。了解线粒体的遗传规律可以更好地认识线粒体疾病的病因学与发病机制。

（一）mtDNA 具有半自主性

线粒体具有自己的遗传物质，所以有人将 mtDNA 称为第 25 号染色体或 M 染色体。mtDNA 能够独立地复制、转录和翻译，但是大量的维持线粒体结构和功能的大分子复合物，以及大多数氧化磷酸化酶的蛋白质亚单位均由核 DNA 编码，故 mtDNA 的功能又受核 DNA 的影响，因而是一种半自主复制体。

（二）线粒体基因组所用的遗传密码和通用密码不同

线粒体的遗传密码和通用密码并不完全一样，核基因组中 UGA 为终止密码子，而在 mtDNA 中 UGA 编码色氨酸。另外，线粒体的 tRNA 兼并性较强，仅用 22 个 tRNA 就可识别多达 48 个密码子。

（三）mtDNA 为母系遗传

人类受精卵中的线粒体来自于卵细胞，也就是说来自母系，这种传递方式称为母系遗传（maternal inheritance）。这是因为卵细胞含有十多万个 mtDNA 分子，而精子只有大约几百个，相对于卵子而言，精子对线粒体基因型的影响很小。另一原因是用于推动精子运动的大量线粒体存在于精子基底部，

在受精时精子尾部会丢失，从而导致精子中的 mtDNA 不能进入卵细胞。由于受精过程和细胞分裂过程中线粒体与细胞核的行为不同，导致线粒体遗传病的传递模式不同于经典的孟德尔遗传。因此，如果在某个家族中发现一些成员具有相同的临床症状，并且是从受累的女性传递下来，就应考虑线粒体 DNA 突变的可能性。

（四）mtDNA 在细胞分裂过程中的复制分离与遗传瓶颈现象

线粒体 DNA 在减数分裂和有丝分裂期间都要经过复制分离（replicative segregation）。人类卵母细胞中虽含有约十多万个线粒体，但在卵母细胞成熟中绝大多数线粒体会丧失，数目减至 10～100 个。这种卵细胞成熟过程中线粒体数目从十万个锐减到少于 100 个的过程就是遗传瓶颈（genetic bottleneck），使得只有少数线粒体真正传给后代，也是造成亲代与子代之间差异的原因。此后，经过早期胚胎细胞分裂，线粒体通过自我复制使数目达到每个细胞含有十万个或更多。如果通过遗传瓶颈保留下来的一个线粒体碰巧携带一种突变基因，那么这个突变基因就可能在发育完成之后的个体中占有一定的比例。由于在胚胎发生和组织形成的细胞分裂过程中线粒体经过复制分离，随机进入子细胞。因此，一些子细胞很可能接受大量的携带突变基因的线粒体，由它们形成的成体组织细胞会具有较高比例携带突变基因的线粒体。

（五）mtDNA 的异质性与阈值效应

如前所述，人类每个细胞中都有数千个乃至十万个 mtDNA 分子。纯质（homoplasmy）是指一个细胞或组织中所有的线粒体具有相同的基因组，即都是野生型序列，或者都是突变型序列。异质（heteroplasmy）则表示一个细胞或组织既含有野生型，又含有突变型线粒体基因组。在异质性细胞中，突变型与野生型 mtDNA 的比例决定了细胞是否出现能量短缺。如果携带突变型线粒体比例较小，则产能不会受到明显影响。相反，当含有大量突变型线粒体基因组的组织细胞所产生的能量不足以维持细胞的正常功能时，就会造成组织中能量供应水平降低，进而影响组织的功能并出现异常的性状。也就是说，当突变的 mtDNA 达到一定的比例时，才有受损的表型出现，这就是阈值效应（threshold effect）。这种线粒体基因突变产生有害影响的阈值效应明显依赖于受累细胞或组织对能量的需求。因此，高需能的组织，如脑、骨骼肌、心脏和肝脏等，更容易受到 mtDNA 突变的影响。

（六）mtDNA 的突变率极高

mtDNA 的突变率比核 DNA 高 10～20 倍。这种高突变率造成个体及群体中 mtDNA 序列差异较大。比较任何两个人的 mtDNA，平均每1000 个碱基对中就有4 个不同。人群中含有多种中性到中度有害的 mtDNA 突变，且高度有害的 mtDNA 突变也会不断增多。不过有害的突变会由于选择而被消除，故突变的 mtDNA 基因虽然很普遍，但线粒体遗传病却不常见。

第二节　线粒体基因突变与常见线粒体遗传病

1962 年，Lufe 发现一位年轻的基础代谢率异常增高的瑞典妇女，同时伴有线粒体结构异常和氧化磷酸化功能异常。这是人类首次认识到线粒体与人类疾病的发生有关。1987 年，Wallace 报道了首例由线粒体 DNA 突变引起的人类疾病，明确了 mtDNA 突变可引起人类疾病。其后的十多年中，这一研究领域进展迅速，现已发现 50 多种 mtDNA 点突变和 100 多种 mtDNA 重排与人类多系统紊乱相关。

当某一组织内的突变型线粒体与野生型线粒体的比例达到一定阈值时，就会产生病变。这个阈值因人而异、因组织而异，与组织消耗的能量、氧，代谢需要的物质及线粒体特定突变引起的效应有关。线粒体突变所表现出的一些临床特征包括：肌病、心肌病、痴呆、突发性肌阵挛、耳聋、失明、贫血、糖尿病和大脑供血异常（休克）。这些临床缺陷的形成与严重程度依赖于多种因素，例如，胚胎发育早期线粒体突变基因的复制分离程度、突变的线粒体基因在某一特定组织中存在的

数量，以及组织中突变的线粒体 DNA 所需达到的阈值水平等等。因此，确定是否存在线粒体基因突变是一个非常复杂的过程。线粒体基因突变的母系遗传特性确定了线粒体疾病的母系遗传。然而，由于某些突变的线粒体基因不能够通过遗传瓶颈，所以并不是所有线粒体病都遵循母系遗传方式。

一、线粒体基因突变的类型

（一）碱基突变

1. 错义突变　通常发生于 mtDNA 中的蛋白质编码序列上，导致所编码的氨基酸发生改变。主要与脑脊髓性及神经性疾病有关，如 Leber 遗传性视神经病（Leber hereditary optic neuropathy，LHON）和神经肌病等。

2. 蛋白质生物合成基因突变　多为 tRNA 基因突变。与错义突变引起的疾病相比，这类突变所致的疾病更具系统性的临床特征，而且几乎所有突变都是 tRNA 突变，并与线粒体肌病相关。典型的疾病包括肌阵挛性癫痫伴碎红纤维病（myoclonnus epilepsy and ragged-red fibers，MERRF）综合征、线粒体肌病脑病伴乳酸酸中毒及中风样发作（mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes，MELAS）综合征、母系遗传的肌病及心肌病等。

（二）缺失、插入突变

以缺失突变更多见。这类疾病往往无家族史，散发。导致 mtDNA 缺失的原因多为 mtDNA 的异常重组或在复制过程中的异常滑动。常见于神经性疾病及一些退化性疾病，如 KSS 综合征（Kearns-Sayre syndrome，KSS）。绝大多数的眼肌病是由缺失突变引起。

（三）mtDNA 拷贝数目突变

拷贝数目突变指 mtDNA 拷贝数大大低于正常，这种突变较少，仅见于一些致死性婴儿呼吸障碍、乳酸中毒或肌肉病变及肝、肾衰竭的病例。

此外，mtDNA 病变还具有相应的组织特异性。不同组织对氧化磷酸化依赖性的差异是线粒体病组织特异性的基础。有人认为，这种依赖性的差异是由核 DNA 编码的氧化磷酸化基因的组织特异性调控造成的。应注意的是，氧化磷酸化过程中 5 种酶复合物是由 mtDNA 和核 DNA 共同编码，编码这些酶的核基因突变也可能产生类似于线粒体病的症状。因此，有些线粒体遗传病是核 DNA 与mtDNA 共同作用的结果。

二、常见线粒体遗传病

作为细胞的能量代谢中心，线粒体一旦出现功能改变就会导致病理变化。随着对线粒体生物化学和遗传学认识的不断深入，所发现的线粒体遗传病也在逐渐增多。人类首先识别的线粒体疾病是 Leber 遗传性视神经病。致病原因为电子呼吸链酶复合体 I 中的亚单位 NADHQ 氧化还原酶基因发生突变所致，其临床表现为在中年时突发失明。部分线粒体基因突变可破坏其对 NADH 的利用能力，另一部分突变则能够阻断电子传递给辅酶 Q。数十年的线粒体基因突变的累积则会导致生物个体衰老、退行性疾病和肿瘤的发生。

由于线粒体是母系遗传，而且卵细胞线粒体的数目非常多，线粒体突变并非涉及所有的线粒体，这也是线粒体疾病复杂的病理表型的分子机制。在一个线粒体疾病家族中，由于突变型线粒体在线粒体总数中所占比例不同，家族成员的临床表型可以从正常表型到非常严重的综合征并存，并且患者的发病年龄也不尽相同。只有细胞中突变型线粒体达到一定比例，线粒体产生能量的能力下降到一定的阈值时，细胞才会丧失其正常的功能。高度依赖于氧化磷酸化的高需能组织器官，例如，神经系统和心脏，在 mtDNA 发生突变时遭受的损害更为严重。

线粒体疾病种类很多，而且原因各不相同。一部分疾病完全是由于 mtDNA 的异常引起的，如 Leber 遗传性视神经病、MERRF 综合征等；还有一部分疾病，发病原因可能部分与线粒体异常有关，如

糖尿病、癌症及心肌疾病的产生、乳酸酸中毒、某些肌病、骨质疏松症、阿尔茨海默症、帕金森病、中风等。此外，人类的衰老也与 mtDNA 突变有关。

(一) Leber 遗传性视神经病(MIM 535000)

Leber 遗传性视神经病或遗传性视神经萎缩最早由德国眼科医师 Leber 发现，是一种母系遗传的视网膜神经节细胞和轴突退化性疾病，这种退化导致急性或亚急性的中心视觉丧失，患者多为成年男性。典型的 LHON 首发症状为视物模糊，随后的几个月之内出现无痛性、完全或接近完全的失明。通常是两眼同时受累，或在一只眼睛失明不久，另一只也很快失明。视神经和视网膜神经元的退化是 LHON 的主要病理特征。另外还有周围神经的退化、震颤、心脏传导阻滞和肌张力降低。LHON 从儿童时期一直到 70 多岁都可能发病，但通常在 20 ~ 30 岁时发病。

在 9 种编码线粒体蛋白的基因(*ND*1，*ND*2，*CO*1，*ATP*6，*CO*3，*ND*4，*ND*5，*ND*6 和 *CYTB*)中，至少有 18 种错义突变可直接或间接地导致 LHON 表型的出现。LHON 分为 2 种类型：①单个线粒体突变就足以导致 LHON 表型；②少见的、需要二次突变或其他变异才能产生的临床表型，但其发病的生物学基础尚不完全清楚。对于第一种类型的 LHON 而言，90% 以上的病例中存在 3 种突变(*MTND*1* *LHON*3460A，*MTND*4* *LHON*11778A，*MTND*6* *LHON*14484C)，而且在这些患者中，11778A 突变占 50% ~80%。在这类 LHON 家族中，常见 mtDNA 纯质性。在异质性 LHON 家族中，突变线粒体 DNA 的阈值水平≥70%。

11778A 突变使电子呼吸链酶复合体 Ⅰ 中的亚单位(NADH 脱氢酶)上第 340 位 G 突变为 A，使高度保守的精氨酸替换为组氨酸，降低了 NAD 关联底物的氧化作用效率。3460A 突变降低了复合物 Ⅰ 近 80% 的活性，14484C 突变也降低了复合物 Ⅰ 的活性，复合物 Ⅰ 在光诱导的神经传导通路中具有非常重要的作用。这三种主要的 LHON 突变都不同程度地影响了呼吸链的作用。

LHON 的致病性突变会影响线粒体氧化磷酸化作用和产生 ATP 的能力，最主要的受累对象是那些高度依赖氧化磷酸化的组织。因此，线粒体成分的缺陷只对某些特定组织产生影响，而不会出现多组织、多器官受累的综合征表型。中枢神经系统(包括脑和视神经)对氧化代谢的需求非常高，这与 mtDNA 突变导致 LHON 的首发临床表现为失明相一致。

(二) MERRF 综合征(MIM 545000)

MERRF 综合征即肌阵挛性癫痫伴碎红纤维病，是一种罕见的、异质性线粒体肌病，有明显的母系遗传特点，具有多系统紊乱的症状，包括肌阵挛性癫痫的短暂发作、不能够协调肌肉运动(共济失调)、肌细胞减少(肌病)、轻度痴呆、耳聋、脊髓神经的退化等等。碎红纤维(ragged-red fibers)是指大量的团块状异常线粒体聚集在肌细胞中，电子传导链中复合物 Ⅱ 的特异性染料能将其染成红色。一般来讲，MERRF 是线粒体脑肌病的一种，包括线粒体缺陷和大脑与肌肉功能的变化。在严重患有 MERRF 综合征患者的大脑卵圆核和齿状核中发现神经元的缺失，并且在小脑、脑干和脊髓等部位也可观察到上述现象。MERRF 病一般在童年时初次发病，病情可持续若干年。

大部分 MERRF 病例是线粒体基因组的 $tRNA^{Lys}$ 基因点突变的结果(A8 344G)。这个突变正式的名称为 *MTTK* MERRF*8344G。线粒体碱基替换疾病的命名包括三个部分：第一部分是确定的位点，MTTK 中的 MT 表示线粒体基因突变，第二个 T 代表 *tRNA* 基因，K 表示赖氨酸，这说明突变发生在线粒体的 $tRNA^{Lys}$ 基因上。第二部分是在星号之后使用了描述临床特征的疾病字母缩略词，这些临床特征与特定核苷酸位点的碱基突变密切相关，缩略词就是 MERRF。第三部分中的术语 8344G 表示在核苷酸 8344 位置的鸟嘌呤(G)的变异。

如果神经和肌肉细胞中 90% 的线粒体存在 *MTTK* MERRF*8344G 突变，就会出现典型的 MERRF 症状，当突变的线粒体所占比例较少时，MERRF 的特征也随之减轻。这种 MERRF 突变减少了线粒体蛋白的整体合成水平，产生了一系列 MERRF 特定的翻译产物，而且除了复合物 Ⅱ，所有氧化磷酸化成分的含量均降低。

（三）MELAS 综合征（MIM 540000）

MELAS 综合征又称线粒体肌病脑病伴乳酸酸中毒及中风样发作综合征，是最常见的母系遗传线粒体疾病的一种，在 1984 年由 Pavlakis 首次发现并报道。临床特点包括：40 岁以前开始的复发性休克、肌病、共济失调、肌阵挛、痴呆和耳聋。少数患者出现反复呕吐，周期性的偏头痛，糖尿病，眼外肌无力或麻痹，从而使眼的水平运动受限（进行性眼外肌麻痹，PEO），并出现眼睑下垂、肌无力、身材矮小等。乳酸性酸中毒是由于乳酸浓度的增高导致血液 pH 值下降和缓冲能力降低。在 MELAS 患者中，异常的线粒体不能够代谢丙酮酸，导致大量丙酮酸生成乳酸，使后者在血液和体液中累积。MELAS 患者的一个特征性病理改变就是在脑和肌肉的小动脉和毛细血管管壁中出现大量形态异常的线粒体聚集。MELAS 虽与 MERRF 的症状相似，但有其独特的临床表现。

在 MELAS 病例中，*MTTL1* * *MELAS* 3243G 突变的发生率超过了 80%，碱基突变发生在两个 $tRNA^{leu}$ 基因之一。发生在 $tRNA^{leu(UUR)}$ 基因上的 A3243G 突变中，UUR 代表亮氨酸 tRNA 的密码子，前两个位置是尿嘧啶，第三个位置（R）为嘌呤。一般情况下，*MTTL1* * *MELAS*3243G 是异质性的，当肌肉组织中突变的 mtDNA ≥90% 时，复发性休克、痴呆、癫痫和共济失调的发病风险就会增加。当 A3243G 突变的异质性达到 40% ~50% 的时候，就有可能出现慢性进行性眼外肌麻痹（CEPO）、肌病和耳聋。此外，MELAS 基因突变还可发生在 $tRNA^{leu(UUR)}$ 基因内 3252、3271 和 3291 位点上，以及线粒体 $tRNA^{Val}$（*MTTV*）与 COXⅢ（*MTCO3*）基因上。

不同种类线粒体突变所导致的临床表现是复杂多变的。除了 MELAS，*MTTL1* 基因中的各种单核苷酸突变也能够产生复杂多变的线粒体遗传病表型。在一些有 A3243G 突变的个体中，唯一的表型是糖尿病和耳聋；而在 3250、3251、3302、3303 和 3260 位点突变的患者中，肌病是其主要特点；心肌病则是 3260 和 3303 位点碱基替换患者的主要症状；存在 C3256T 突变的患者则表现出 MELAS 和 MERRF 两种疾病的共同症状。总而言之，不同的线粒体 *tRNA* 基因突变可引起不同的功能紊乱，一些线粒体 *tRNA* 基因突变能产生相似的临床症状，而同一 *tRNA* 基因不同位点的突变又能导致不同的临床表型。

（四）KSS 病（MIM 530000）

KSS 病又称为慢性进行性眼外肌麻痹，因进行性眼外肌麻痹和视网膜色素变性而得名。KSS 的表现还包括心肌电传导异常、小脑性共济失调、近端肌无力、耳聋、糖尿病、生长激素缺乏、甲状旁腺功能减退症或其他的内分泌病。在 KSS 病中，肌肉病变最初是单侧的，但是通常发展为双侧病变，病程呈进展性，发病年龄一般低于 20 岁，大多数患者在确诊后几年内死亡。

KSS 并不表现出特定的母系或核基因遗传方式，但其症状表明它仍是一种线粒体疾病。KSS 患者的线粒体 DNA 存在结构上的改变，包括大片段缺失（>1000bp）和 DNA 复制。线粒体基因组的这种异常可以通过 Southern 杂交检测，使用线粒体特异性 DNA 探针可以确认受累者线粒体中存在的复制或缺失，而后借助序列分析确定 mtDNA 结构异常的性质和程度。大约 1/3 的 KSS 病例与 4977bp 缺失有关，该缺失的断裂点位于 *ATP8* 和 *ND5* 基因内，并伴随间隔结构和编码 *tRNA* 基因的缺失。大多数的 KSS 病例是散发的，但不排除由无症状的母亲遗传而来的可能性。

KSS 的病情严重性是由异质性的程度和 DNA 结构发生改变的线粒体基因的组织分布决定。当肌细胞中突变的线粒体基因比例多于 85% 时，可出现 KSS 所有的临床特征。在异质性处于较低水平时，主要症状是进行性眼外肌麻痹。当缺失或（和）复制的线粒体基因在造血干细胞中大量存在时，就会表现出一种致命且早发的疾病，称 Pearson 综合征（Pearson syndrome，PS）。PS 的主要特点是血细胞不能利用铁来进行血红蛋白的合成，从而引起缺铁性贫血。

当存在缺失的线粒体 DNA 分子在某一组织中的含量非常高时，由于线粒体部分 DNA 包括 tRNA

基因的丧失，能量的产生就会急剧下降。同样，当含有复制的线粒体基因增加时，线粒体基因（包括 tRNA 基因）的过度表达将会导致氧化磷酸化（OXPHOS）亚基的失衡，从而影响呼吸链中蛋白复合物的组装。

Summary

Mitochondria, which are often called 'oxidation central and energy plants' or 'powerhouses', are essential for all eukaryotic cells for sustaining life because of the oxidative respiration that takes place in mitochondria to generate the adenosine triphosphate. As the only organelle containing extranuclear DNA in eukaryotic cell, mitochondrion exists in all types of cells (except red blood cells) and tissues. Human mtDNA consists of 16,569 base pairs (bp), which exists as a double-stranded circular molecule (Fig.7-1). The outer stand is named as heavy strand (H) while the inner strand is light strand (L). Each mtDNA contains genes encoding of 13 proteins (involved in electron transport chain in mitochondria), 22 tRNAs and 2 rRNAs. Compared to nuclear DNA, mtDNA has very unique inheritance pattern and characteristics which differentiate itself from nuclear DNA including semiautonomous replication, different genetic codes from the universal codes, maternal inheritance, replication and segregation during both mitosis and meiosis, threshold effect, and high mutation rate. It is known that mtDNA mutations are involved in human diseases. Over 50 mononucleotide mutations in mitochondrial genomes have been shown to be associated with numerous dysfunctions of the organs. The clinical manifestations caused by mtDNA mutations include myopathy, cardiomyopathy, dementia, explode myoclonus, deafness, acroisa, anemia, diabetes and cerebrovascular dysfunction. The degree of these clinical defects depend on many factors, such as the number of mutant mtDNA molecules in a particular tissue and the minimum number of mutant mtDNA molecules required to cause mitochondrial dysfunction in the particular tissue. Therefore, the diagnosis of a mitochondrial disease requires a study of the family history and sequencing analysis of mtDNA molecules. Although the majority of mitochondrial diseases is maternally inherited mutations in mtDNA, somatic mutation of mtDNA may explain some of the cases.

（傅松滨）

参考文献

1. http://www.mitomap.org
2. http://en.wikipedia.org
3. BrownMD, Voljavec AS, Lott MT, et al. Leber's hereditary optic neuropathy: a model for mitochondrial neurodegenerative diseases. FASEB Journal, 1992, 6(10): 2791-2799
4. DiMauro S, Hirano M. Mitochondrial DNA Deletion Syndromes. GeneReviews. April 19, 2007, 13: 2009
5. Filosto M, Tomelleri G, Tonin P, et al. "Neuropathology of mitochondrial diseases". Biosci. Rep, 2007, 27(1-3): 23-30
6. Gardner A, Boles RG. "Is a "Mitochondrial Psychiatry" in the Future? A Review". Curr. Psychiatry Review, 2005, 1(3): 255-271
7. Hirata K, Akita Y, Povalko N, et al. "Effect of l-arginine on synaptosomal mitochondrial function". Brain Dev, 2007, 30: 238
8. Henze K, Martin W(2003). "Evolutionary biology: essence of mitochondria". Nature, 2003, 426(6963): 127-128
9. Koga Y, Akita Y, Nishioka J, et al. "MELAS and L-arginine therapy". Mitochondrion, 2007, 7(1-2): 133-1399
10. MayorovV, BiousseV, Newman N J, et al. The role of the ND5 gene in LHON: Characterization of a new, heteroplasmic

LHON mutation. Annals of Neurology 58(5):807-811

11. NINDS Kearns-Sayre *Syndrome Information* Page. National Institute of Neurological Disorders and Stroke(NINDS). February 13,2007. Accessed January 13,2009
12. ValentinoML,Barboni P,Ghelli A,et al. The ND1 gene of complex I is a mutational hot spot for Leber's hereditary optic neuropathy. Annals of Neurology 56(5):631-641
13. Yu-Wai-ManP, GriffithsPG, Hudson G, et al. Inherited mitochondrial optic neuropathies. J Med Genet, 2009, 46: 145-158
14. ZhadanovSI, AtamanovVV, Zhadanov N I, et al.. A novel mtDNA ND6 gene mutation associated with LHON in a Caucasian family. Biochemical and Biophysical Research Communications, 2005, 332(4):1115-1121

第十章 药物反应的遗传基础

临床医师都知道不同的患者对同一药物有不同的反应。举个古老的例子，Hanzlik（1913）曾报道对水杨酸钠的不良反应有个体差异。他调查300例服用水杨酸钠的男性病员，约有2/3病员在水杨酸钠总量达65～130g时发生不良反应，但少数敏感病员在水杨酸钠总量仅为3.25g时即出现不良反应，相反有个别耐受病员要在水杨酸钠总量达到130.0g左右时才有不良反应。可见引起不良反应的水杨酸钠剂量在不同个体间相差几乎达40倍。这种例子举不胜举。药物在体内的作用，取决于机体对它的吸收，它在器官间的分布，它与细胞受体的相互作用以及它的代谢过程。这些过程必定要受环境因素如食物和其他药物的影响，但无疑也要受遗传因素的控制。1959年，Vogel提出了药物遗传学（pharmacogenetics）这一术语，主要从单个基因的角度对药物与遗传之间的关系进行研究，认为药物代谢过程中涉及各种酶和受体，它们都是在基因控制下合成的蛋白质。如果基因突变产生了异常的酶或由于酶的缺乏，产生了异常的蛋白质，形成了异常的药物受体，药物代谢过程就要发生改变，从而引起异常的药物反应。到了20世纪90年代，随着人类基因组计划的提出和不断实施，一大批人类基因相继被发现，人们对药物与遗传之间关系的研究也在不断的深化。药物在人体内的代谢是一个十分复杂的过程，与遗传之间的关系也非常复杂，不能用单个基因的方法来进行研究，而是要把它放在基因组这个整体中加以考虑。因此药物基因组学（pharmacogenomics）这个概念也就在1997年应运而生了。药物基因组学以药物效应安全性为目标，研究各种基因突变与药效及安全性之间的关系。

第一节 药物遗传学

一、药物代谢的遗传研究

机体摄入药物后，经过吸收、分布，在细胞内使药物发生生物转化后失去药效，排出体外。这是药物代谢的过程，也叫解毒（detoxi-cation）。药物解毒作用主要在肝脏进行，有些被彻底氧化为二氧化碳后呼出，有些以各种不同形式被排出，或是通过肾脏由尿排出，或是通过肝脏由胆汁随粪便排出。许多药物通过种种生化反应提高了溶解度，使之易于排出。其中一个重要的生化反应是

与葡萄糖醛酸(glucuronic acid)结合,如吗啡(morphine)及可待因(codeine)等吗啡衍生物的排出全依靠这一结合作用。异烟肼(isoniazid)在被乙酰化(acetylation)后排出。磺胺类药物也被乙酰化后排出。

研究一种药物的代谢及其效应,通常是给予一份标准剂量,经适当时间间隔后测定血液中药物水平或其他表示药物代谢速率的参数。药物反应的个体变异可以是连续的或是不连续的。如取一大样本进行测定,将受试者的反应分类后作图,连续变异者可得到一条钟形曲线,呈单众数分布;在不连续的变异,其频数分布曲线将呈双众数,有时为三众数。单众数分布意味着该药物代谢受控于多基因,分析其遗传机理目前将是困难的。双众数或三众数分布意味着该药物受单基因控制。如某一药物的正常代谢受显性基因 R 控制,有些人因为是隐性基因纯合 rr 而对该药物无代谢能力,这时就有三种基因型:RR、Rr、rr。如果纯合子 RR 和杂合子 Rr 的表现型不能区分,即它们对该药物有同样反应,将为双众数分布。如果纯合子 RR 和杂合子 Rr 的表现型可以区分,即它们对该药物的反应略有不同,将为三众数分布,每个峰或每个众数代表一种表现型(见图 10-1)。

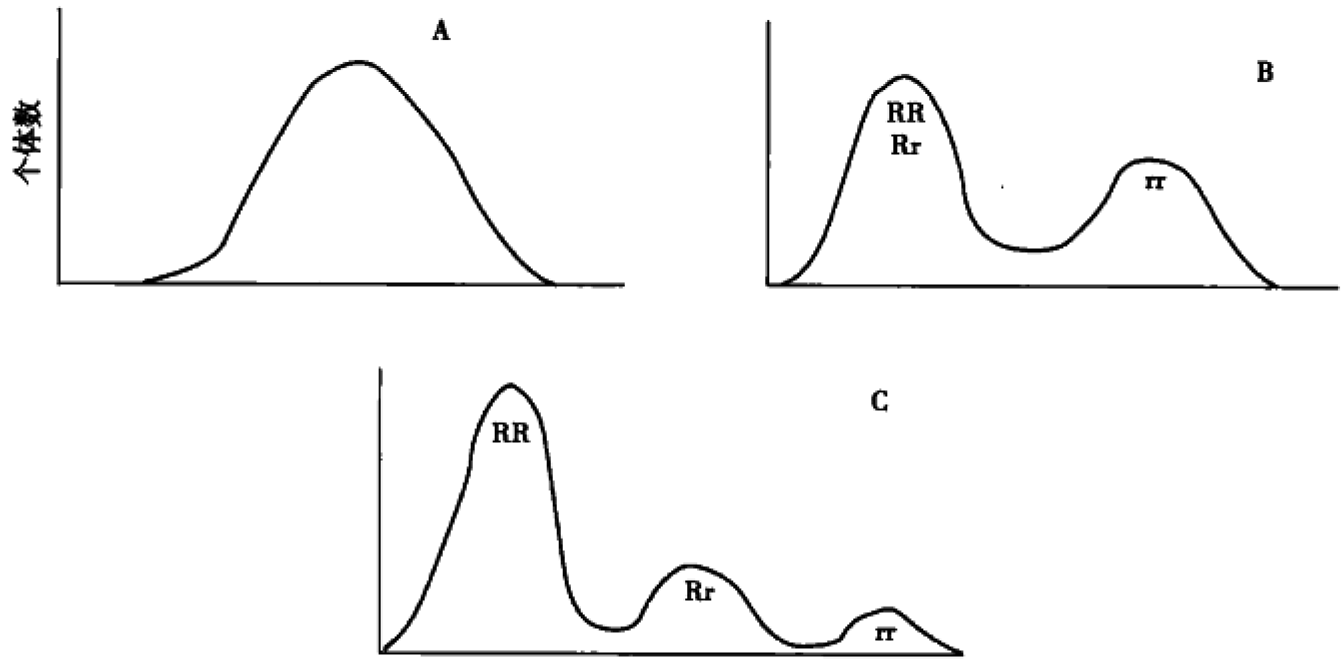


图 10-1 对不同药物的各种反应方式

有些药物反应差异并非遗传决定,如青年人和老年人对吗啡及其衍生物比较敏感。肝病患者也对吗啡及其衍生物较为敏感。药物遗传学研究的是由遗传决定的药物反应差异(表 10-1)。

表 10-1 有异常药物反应的遗传性状

性状或缺乏的酶	受累系统	药物或作用因素	性状频率	临床影响
无过氧化氢酶	各种组织	过氧化氢	日本人群高	对无过氧化氢无反应
α_1-抗胰蛋白酶	血浆	吸烟	比较稀有	肺气肿
乙酰转移酶	肝	异烟肼	常见	多发性神经炎、异烟肼肝炎
血清胆碱酯酶缺乏	血浆	琥珀酰胆碱	比较稀有	呼吸暂停
G-6-PD 缺乏	红细胞	蚕豆、伯氨喹啉等	地中海地区及黑人中高,我国华南多见	药物性溶血
双羟香豆素抗性	凝血	双羟香豆素	稀有	抗凝作用降低

续表

性状或缺乏的酶	受累系统	药物或作用因素	性状频率	临床影响
卟啉症	肝	巴比妥酸盐、磺胺类等	多变	急性发作
青光眼	眼	糖皮质激素	常见	眼压升高
周期性麻痹	？细胞膜	胰岛素、肾上腺素等	稀有	麻痹
高铁血红蛋白还原酶缺乏	红细胞	氨苯砜、伯氨喹啉等	多变	先天性高铁血红蛋白血症
不稳定血红蛋白	红细胞	磺胺类、氧化剂	稀有	溶血
恶性高热	？肌肉神经	麻醉剂	稀有	寒战高热
Parkinson 综合征	脑	左旋多巴	稀有	舞蹈样动作

二、由药物代谢异常揭示的遗传变异

（一）无过氧化氢酶症

无过氧化氢酶症（acatalasia）患者不能分解过氧化氢，此症系由日本耳鼻喉科医师 Takahara 于 1959 年发现。他在为一个 8 岁女孩治疗口腔坏疽时，为了清除坏疽组织，用过氧化氢溶液消毒。在正常情况下由伤口渗出的血液是鲜红的，并有气泡发生。但这一女孩的伤口渗血与过氧化氢接触后转为棕黑色，而且不出现气泡。Takahara 提示此女孩可能缺乏能将过氧化氢分解为水和氧的过氧化氢酶（catalase）。后来的研究证明确为过氧化氢酶缺乏，患者不能分解过氧化氢，伤口渗血中的血红蛋白被氧化为棕黑色的高铁血红蛋白。家系分析表明本症系常染色体隐性遗传。测定血清过氧化氢酶水平可区别三种表现型：正常基因纯合子有正常酶活性，无过氧化氢酶基因纯合子血液中无酶活性，杂合子有中等水平酶活性。本症不限于日本，在世界其他地区也有发现。现已发现有五种不同类型，其病情从轻（牙槽溃疡）到重（齿根坏疽和牙槽萎缩）。受累者仅约半数有口腔疾患，另外一半人是完全健康的。曾报道瑞士有一种很稀有的无过氧化氢酶症。瑞士型少量残存的酶的理化特性与正常酶不同，日本型残存的酶是正常的。这提示瑞士型可能代表结构基因突变而日本型则系由于控制基因突变。过氧化氢酶基因位于 11p13。我国此症的发生率在 1% 以下，其中华北的发生率为 0.65%，华中为 0.55%，华南为 0.23%，台湾为 0.29%。

（二）α_1-抗胰蛋白酶缺乏

α_1-抗胰蛋白酶（α_1-antitrypsin）能抑制胰酶及其他一些酶的蛋白水解作用。突变基因 Pi（表示蛋白酶抑制物）纯合子缺乏这种酶，其频率约为 1/1500。已检出好几个等位基因。很多患有肺气肿的人为这一突变基因纯合子，属常染色体隐性遗传。推断吸烟等原因引起炎症而导致放出胰蛋白酶。在正常情况下，胰蛋白酶为 α_1-抗胰蛋白酶所抑制；但在 α_1-抗胰蛋白酶缺乏者，胰蛋白酶则损害肺组织。杂合子如果吸烟的话也易得梗死性肺部疾患。杂合子约占人群 5%。如果把吸烟也看作是一种服药的话，这也可视为是一种药物遗传病。

（三）异烟肼快失活者和慢失活者

异烟肼（isoniazid）是治疗结核病最重要的药物之一。此药吸收迅速，在体内失活后排出。有些患者排出很快，有些患者排出很慢，分别称为快失活者（rapid inactivator）和慢失活者（slow inactivator）。家系分析表明，快失活者系按常染色体显性性状传递，慢失活者为隐性基因纯合子，杂合子使异烟肼失活的速度居中。涉及的酶是肝脏中的乙酰转移酶（acetyltransferase）。欧美人群中约 50% 为慢失活者，东方人群约低于 20%，我国情况有待测定。

有些人服用异烟肼可引起多神经炎（polyneuritis）。这是因为异烟肼可与吡哆醇（pyridoxine）

(维生素 B_6)相互作用,导致维生素 B_6 缺乏。当然这只发生于慢失活者,因为他们血液中的异烟肼浓度在较长时间内维持较高水平。给予维生素 B_6 可预防这一并发症。在必须间歇给药时,快失活者治疗反应较差。慢失活者对治疗结核病并不更为有利。异烟肼失活的表现型看来不影响治疗反应。

抗抑郁剂苯乙肼 (phenelzine) 的分子构型与异烟肼相似,也在体内被乙酰化。并非所有患者都对苯乙肼有治疗反应,根据患者的异烟肼失活能力可预期哪些患者有治疗反应。最近研究提示异烟肼慢失活者对苯乙肼的治疗反应较异烟肼快失活者为佳,但慢失活者会有视力模糊等副反应。

对同时患有结核病和癫痫的患者有一种奇特的遗传性药物相互作用。在用苯妥英钠(phenytoin, Dilantin)治疗癫痫时,体内异烟肼含量高时可抑制肝脏氧化酶对苯妥英钠的作用而使苯妥英钠达到毒性水平。当然这种情况也只发生于异烟肼慢失活者。

异烟肼快失活者有可能发生异烟肼肝炎。Mitchell 等(1976)认为异烟肼在肝内乙酰化后被水解为异烟酸(isonicotinic acid)和乙酰肼(acetylhydrazine)。后者在肝内又形成乙酰作用物质,进而导致肝组织坏死。快失活者肝内有较多的乙酰异烟肼,从而产生较多的乙酰肼而导致异烟肼肝炎。

(四) 伯氨喹敏感

有些人对抗疟药伯氨喹(primaquine)是敏感的。在服药开头几天并无反应,继之则突然尿液变为暗黑色,出现黄疸。随着红细胞被破坏,红细胞计数和血红蛋白浓度下降。患者通常会逐渐恢复,但偶然由于大量溶血而致命。后来发现受累者对伯氨喹敏感是由于缺乏葡萄糖-6-磷酸脱氢酶(glucose-6-phosphate dehydrogenase, G-6-PD) G-6-PD 涉及红细胞酵解过程中的旁路代谢,借以维持红细胞还原型谷胱甘肽(reduced glutathione)的浓度,这是维持红细胞稳定所必需的。G-6-PD缺乏或活性过低时,导致还原型谷胱甘肽缺乏,使红细胞易被破坏而导致溶血。患者红细胞内可见 Heinz 小体(图 10-2)。受累者在进食蚕豆或蚕豆制品时亦可引发溶血危象,故又称为蚕豆病(favism)。

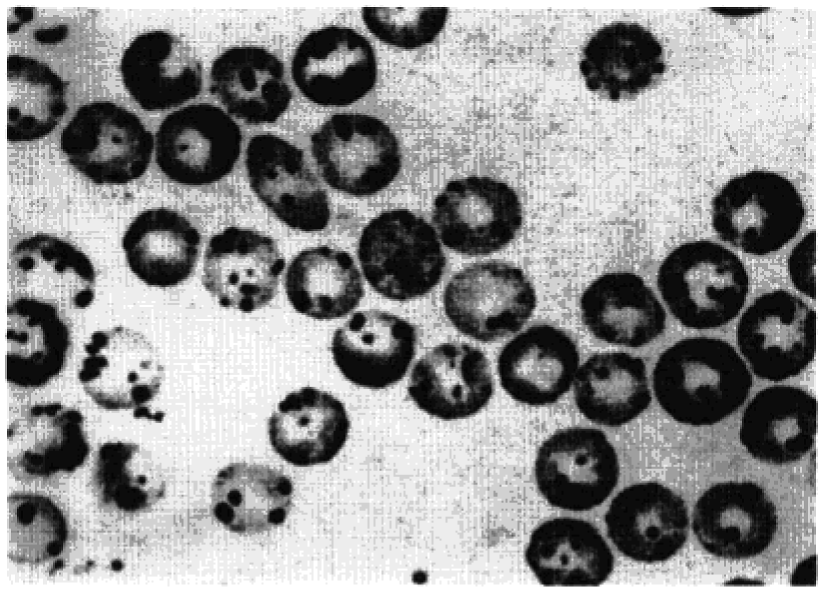


图 10-2 G-6-PD 缺乏者红细胞示 Heinz 小体

G-6-PD 缺乏者不仅对伯氨喹敏感,对其他很多药物也是敏感的,参见表 10-2。

G-6-PD 缺乏是 X 连锁隐性遗传。女性杂合子有两群体细胞,是 Lyon 假说在人体最早的证据之一。两群细胞平均各为 50%,但可有很大变动,甚至达 99:1。由于 X 染色体失活是随机的,如果造血组织来自 n 个细胞,则全部是有突变基因的 X 染色体失活或全部是有正常基因的 X 染色体失活的机会是$(1/2)^n$。如果全部有正常基因的 X 染色体失活,那个女性就会像带有突变基因的男性一样易感(显示杂合子)。大约有 1/3 女性杂合子,其体内突变细胞足以使她们在接触诱发药物时发生显著溶血。一般说溶血的严重程度与 G-6-PD 缺乏的程度相关。使 G-6-PD 缺乏者发生药物性溶血的那些药物也会使某些异常血红蛋白携带者发生溶血,如 Hb H, Hb Zurich 和 Hb Torino。

G-6-PD 缺乏症呈世界性分布,但比较集中在非洲、地中海沿岸、中近东及东南亚、美洲黑人,中美洲及南美洲某些印第安人。我国主要分布在黄河流域以南各省,尤以广东、广西、贵州、云南、四川等省发生率较高,约 5% ~20%。中国人群中 G-6-PD 基因突变型见表 10-3。

表 10-2 可诱发 G-6-PD 缺乏者发生溶血的药物

1. 抗疟药	4. 磺胺类药物
伯氨喹(primaquine)	氨苯磺胺(sulfanilamide, aminobenzensulfonamide)
扑疟喹啉(pamaquine, plasmaquine)	磺胺吡啶(sulfapyridine)
戊氨喹啉(pentaquine)	磺胺异噁唑(sulfisoxazole)
杀疟喹啉(quinocide)	磺胺甲氧吡嗪(sulfamethoxy pyrazine, salfalene)
奎尼丁(quinidine)	磺胺醋酰(sulfacetamide)
喹宁(quinine)	水杨酸偶氮磺胺吡啶(salicy lazosulfapyridine)
2. 解热镇痛药	5. 砜类药物
阿司匹林(乙酰水杨酸)(aspirin, acetylsalicylic acid)	氨苯砜(dapsone, diaminodi phenylfone)
乙酰苯胺(acetylphenylamine)	硫氧二砜(sulfoxone)
非那西丁(phenacetin, acetophenetidin)	6. 其他
安替比林(antipyrine)	萘(naphthalene)
氨基比林(aminopyrine, pyramidon)	苯肼(phenylhydrazine)
3. 呋喃类药物	水杨酰胺(salicylamide)
呋喃丹啶(furadantin, furantoin)	二巯基丙醇(2,3-mercaptopro panol)
呋喃西林(furacilin, furacin)	氯霉素(chloramphenicol)
呋喃唑酮(furaxone)	维生素 K(vitamin K)
呋喃它酮(furaltadone)	蚕豆素(favisin)

表 10-3 中国人 *G6PD* 基因突变型

类型	碱基突变	氨基酸改变	发生的地区
C1	1376 G→T	459 精→亮	中国内地、中国台湾省
C2	1388 G→A	463 精→组	中国内地、中国台湾省
C3	1311 C→T	无	中国内地、中国台湾省
C4	392 G→T	131 甘→缬	中国内地、中国台湾省
C5	1024 C→T	341 亮→苯丙	中国内地、中国台湾省
C6	95 A→G	32 精→组	中国内地、中国台湾省
C7	592 C→T	198 精→半胱	中国内地、中国台湾省
CT1	835 A→T	297 苏→丝	美籍华人
CT2	1360 C→T	454 精→半胱	中国内地、中国台湾省
CT3	493 A→G	165 冬胺→天冬	中国台湾省
	1004 C→A	335 丙→天冬	中国内地

G-6-PD 基因位于 Xq28

第二节 药物基因组学

一、遗传多态性与药物代谢、分布

现在认为,对一个药物的总的药理作用都是由多基因控制的,因为在药物代谢途径中一系列蛋白和酶的基因都会对药物作用产生影响。如以某个药物具有代谢酶多态性与药物作用受体多态性为例,不同的个体就会产生9种不同的药物治疗效果和药物毒性的表现型(图10-3)。在图10-3A中,具有野生型纯合子代谢酶的个体在代谢后留下30%的药物去作用于靶受体,按靶受体野生型和突变型的不同,可产生三种治疗效果和毒性反应。图10-3B和图10-3C中由于代谢基因型是杂合子或突变型纯合子,它们的药物代谢较慢,代谢后留下的药物浓度分别达到65%和99%,它们的治疗效果和毒性也依不同受体基因型而表现出很大差别。

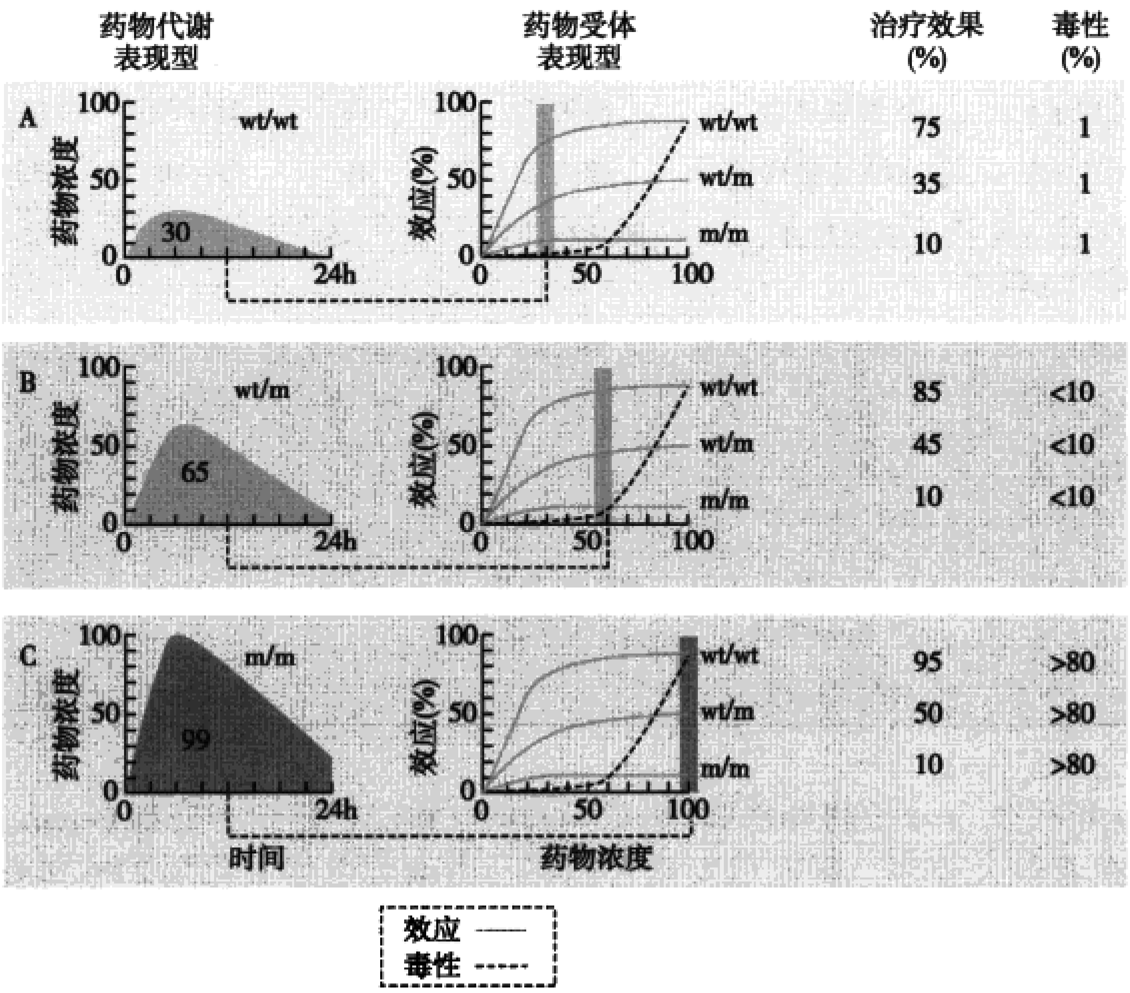


图10-3 药物反应的多基因控制(摘自 Evens and Relling, Science 1999;286:487)

wt/wt:野生型纯合子;wt/m:野生型、突变型杂合子;m/m:突变型纯合子

现以细胞色素氧化酶p450 2D6(CYP2D6)为例,说明遗传多态性与药物代谢的关系。

细胞色素氧化酶p450是存在于肝细胞内质网和线粒体内的氧化酶。主要对药物及其他代谢物进行氧化修饰。细胞色素p450种类很多,有的底物可被几种细胞色素p450催化,而有的细胞色素氧化酶p450可催化几种不同的底物,细胞色素p450的基因命名为*CYP*。

人类的CYP最早是于1984年从快代谢者的肝微粒体中纯化出的,分子量51KD,主要分布在成人肝脏中,肝外分布有脑和胃肠道。脑内的CYP2D6也具有多态性特点,而且影响脑的功能。肝脏中的CYP2D6含量占P450肝脏蛋白总量的1%~2%,虽然含量不高,但参与代谢的药物却占总P450代谢药物的30%。

目前接受的细胞色素氧化酶的命名方法是由CYP后接上阿拉伯数字表述族(1~4)、一个字母表

示亚族(A~F)以及一个阿拉伯数字表明同工酶(1~20)。在星号后面的数字表示等位基因的变异。根据药物代谢重要性递减排列,P450 包括 CYP3A4,CYP2D6,CYP2C9,CYP2C19 和 CYP2A6。它们的主要变异和频率,见表 10-4。

表 10-4 人类细胞色素氧化酶的主要等位基因分布

酶	主要的基因突变	突变种类	酶功能的改变	各人种等位基因频率(%)			
				高加索	亚洲	非洲	依索比亚、沙特
CYP2A6	CYP2A6 * 2	160 亮→组	酶失活	1~3	0	未定	未定
	CYP2A6del	缺失	没有酶产生	1	15	未定	未定
CYP2C9	CYP2C9 * 2	144 精→半胱	对 P450 的黏附性下降	8~13	0	未定	未定
	CYP2C9 * 3	359 异亮→亮	针对底物的特异性改变	6~9	2~3	未定	未定
CYP2C19	CYP2C19 * 2	异常联接点	酶失活	12	23~32	13	14~15
	CYP2C19 * 3	提前终止密码子	酶失活	0	6~10	未定	0~2
CYP2D6	CYP2D6 * 2 × N (N=2,3,4,5,13)	基因重复	酶活性增高	1~5	0~2	2	10~16
	CYP2D6 * 4	缺陷联接	酶失活	12~21			
	CYP2D6 * 5	缺失	没有酶产生	2~7	1	2	1~4
	CYP2D6 * 10	34 脯→丝、486 丝→苏	酶不稳定	1~2	6	4	1~3
	CYP2D6 * 17	107 苏→异亮、296 精→半胱	对底物的黏附性下降	0	51	6	3~9
		486 丝→苏			未定	34	3~9

20 世纪 70 年代发现降压药异喹胍(debrisoquine)和催产药金雀花碱(sparteine)的反应在不同个体中存在着很大差别。研究后发现它们都是由细胞色素氧化酶 2D6 氧化代谢而失活。

大部分个体对以上两种药物代谢正常,称为强代谢者(extensive metabolizer,EM),但有些个体代谢较弱,称为弱代谢者(poor metabolizer,PM),因此药物反应和副作用增加。家系研究显示氧化代谢反应是单基因控制,PM 是隐性遗传。

随着对 *CYP2D6* 多态性研究的不断深入,又发现在人群中除了 PM、EM 外还存在中间代谢型 IM (intermediate metabolism)和极快代谢型 UEM(ultraextensive metabolism)。

人类 *CYP2D* 基因位点定位于 22q13.1,由 *CYP2D8P*,*CYP2D7P* 和 *CYP2D6* 三个基因组成。尽管它们互相之间同源性很高,但 *CYP2D8P* 和 *CYP2D7P* 都有变异,不能转录为特异的 mRNA,因此它们都是假基因。仅 *CYP2D6* 可产生蛋白质产物,*CYP2D6* 基因由 9 个外显子和 8 个内含子组成。由于基因突变导致的 *CYP2D6* 活性降低或缺失见图 10-4。

CYP2D6 表型可以用探药(probe drug)来测定。探药有金雀花碱、异喹胍和右美沙芬。

表 10-5 列举的是三种细胞色素氧化酶在弱代谢者中清除能力降低的主要分子。对于这些酶,除了 CYP2D6 之外,都是在诱导的情况下出现症状。诱导情况与剂量有关,并且都是可逆的。有几种亚族竞争结合同一区域,能产生阻碍作用。

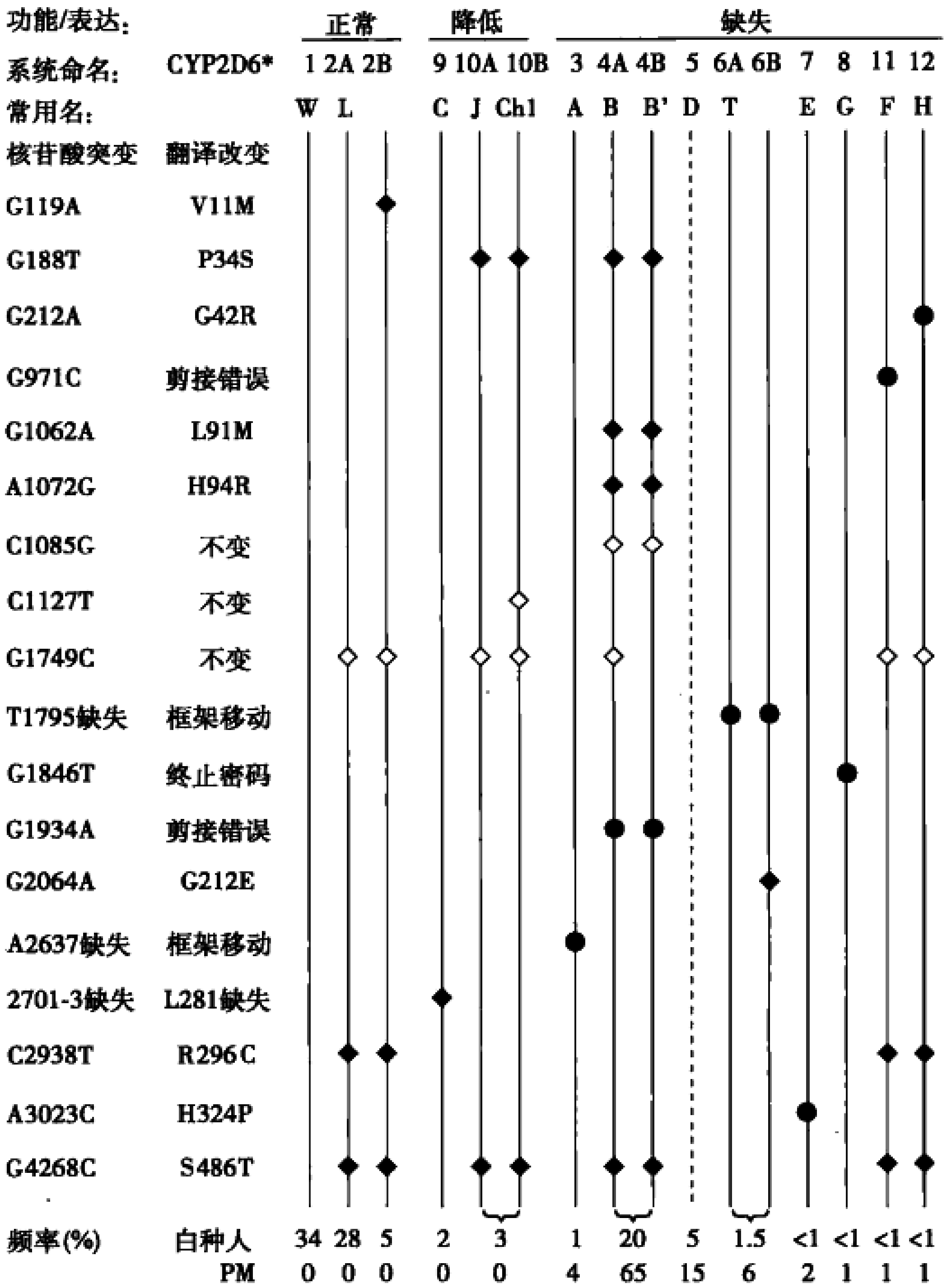


图 10-4　常见 *CYP2D6* 基因突变分布

◆ 为引起氨基酸改变的突变;◇ 为无氨基酸改变的沉默突变; ● 为功能缺失的突变

表 10-5　三种酶的弱代谢者的治疗结果

酶	清除减少的药物	副作用	减少的药物前体形式
CYP2C9	S-华法林	出血	洛沙坦(非肽类血管紧张素Ⅱ的受体拮抗药)
	苯妥英	共济失调	
	洛沙坦		
	甲磺丁脲	低血糖	
	NSAID	消化道出血(?)	
CYP2C19	奥美拉唑	镇定	胍(Proguanil)
	地西泮		
CYP2D6	三环化物抗抑郁药	心脏毒性	曲马朵(镇痛药)
	氟哌啶醇	帕金森(?)	可待因
	抗心律失常药	心律失常	乙基吗啡(Ethylmorphine)
	奋乃静		
	哌克昔林	神经症状	
	5-羟色胺重摄取阻滞剂	恶心	
	珠氯噻醇(抗精神病药)		
	S-米安色林(抗抑郁药)		

鉴定 CYP2D6 表型常用异喹胍代谢率法，即晚上口服异喹胍后收集 8 小时尿样，用气相色谱测定尿中异喹胍及其代谢产物 4-羟基异喹胍含量，然后以下列公式计算异喹胍代谢率(MR)：

MR = 尿中异喹胍量/尿中 4-羟基异喹胍量

MR 大于 12.6 者为 PM，小于 12.6 者为 EM。

PM 个体除对异喹胍-金雀花碱有代谢缺陷外，还对 30 余种重要药物有代谢障碍。如非那西丁，苯乙双胍，恩卡尼，普萘洛尔等。

异喹胍氧化代谢多态性是 CYP2D6 酶的缺陷造成的。通过对弱代谢者肝 mRNA 转录、克隆和测序，目前已鉴定出 *CYP2D6A*、*CYP2D6B*、*CYP2D6C*、*CYP2D6D*、*CYP2D6E*、*CYP2D6F*、*CYP2D6H*、*CYP2D6L*、*CYP2D6T*、*CYP2D6Z* 和 *CYP2D6ch*(或称为 *CYP2D6J*)与异喹胍弱代谢者有关的 *CYP2D6* 突变型等位基因。

CYP2D6 除对药物有灭活作用外，也对有些药物修饰后起激活作用。如可待因就需在 CYP2D6 作用下才能发挥止痛效果，PM 患者对此药有耐受。恩卡尼在 PM 患者中则达不到抗心律失常的作用。PM 发生率在白种人中较高，约 3% ~10%，在黄种人中较低(见表 10-6)。

表 10-6 不同种族 PM 频率

民 族	PM 频率(%)	民 族	PM 频率(%)
白种人	3 ~ 10	西非人	0 ~ 8
中国人	0.7 ~ 1.1	非裔美国人	1.9
日本人	0 ~ 2.3		

除 CYP2D6 外，其他一些细胞色素氧化酶也都呈现多态性，如 CYP2C9，CYP3A4 等。细胞色素氧化酶被认为是药物代谢第Ⅰ期(修饰作用)中的主要酶类。而第Ⅱ期的酶主要起结合反应(图 10-5)。

CYP2D6 代谢多态性与某些疾病的易感性有关，如肿瘤(肺癌、膀胱癌、肝癌、脑瘤等)；帕金森病(Parkinson disease)；阿尔茨海默病(Alzheimer disease)；关节强硬性脊椎炎(ankylosing spondylitics)；

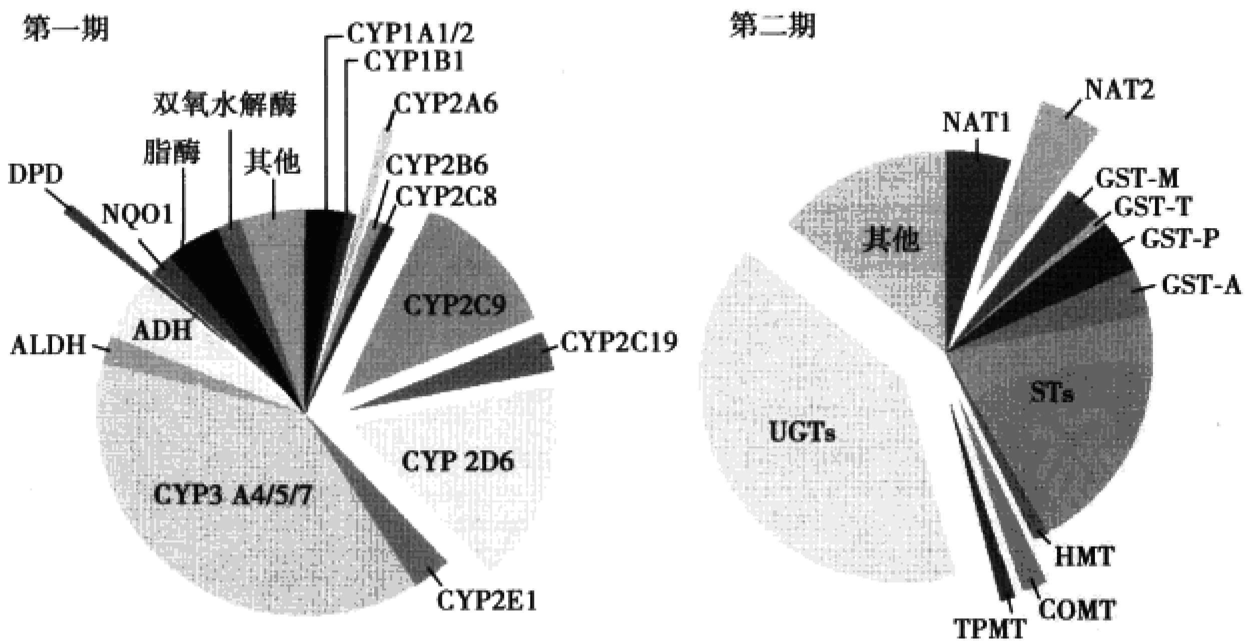


图 10-5 具有多态性的药物代谢酶(摘自 Evens and Relling, Science286:487)

第Ⅰ期的酶主要修饰底物功能基团；第Ⅱ期的酶催化底物的结合反应

ADH：酒精脱氢酶
ALDH：甘油醛脱氢酶
CYP：细胞色素 P450
DPD：二氢尿嘧啶脱氢酶
NQO1：NADPH：泛醌氧化还原酶
COMT：儿茶酚-O-甲基转移酶
GST：谷胱甘肽-S-转移酶
HMT：组胺甲基转移酶
NAT：N-乙酰基转移酶
STs：磺基转移酶
TPMT：巯基嘌呤甲基转移酶
UGTs：三磷酸尿苷葡糖醛酰基转移酶

系统性红斑狼疮(systemic lupus erythematosus)和风湿性关节炎(rheumatoid arthritics)。

细胞色素 P450 的多态性可改变药物的药代动力学,导致药效增强或降低甚至毒副作用的增加,也是产生个体差异的遗传学原因。尤其对于治疗范围窄的药物,药物代谢多态性会产生严重的临床后果。

二、新药开发和临床治疗前景

药物基因组学不仅仅研究与药物代谢有关的基因(这是药物遗传学的主要研究方向),还要研究所有与药物作用有关的基因,这些基因在不同个体中的多样性,以及这些多样性对药物的疗效、毒性或副作用等的影响。因此,药物基因组学的研究将对药物的设计、开发以及应用等产生革命性的改变。

有人预测药物基因组学将会产生重要的新型分子诊断技术,并且会成为临床常规实验室检测技术,医师和药剂师可以参考这些测试结果为患者选择药物和剂量。不同于过去基于经验的逐步解析诊断方法,药物基因组学可以提供每个患者不同的诊断来优化药物治疗方案。从几毫升血液中分离出的 DNA,可以检测几千种基因型,这方面技术发展十分快捷,很快就可以用高输出的自动化系统进行这种检测,并在一次试验中同时检测几千个 SNP。这一点很重要,因为很多药物作用都是多基因作用的结果,所以对于治疗方案的选取也应该基于一批 SNP 的诊断。这些测试结果还要被翻译成患者诊断和治疗选择的临床指标。新的检验手段并不会完全代替传统的治疗,只是为优化用药方案提供新的选择。药物基因组学将使今后的用药和剂量趋于更有效更低毒性的方向发展。

(一) 寻找药物相关基因

每种药物进入体内后都会与许多蛋白质相互作用,例如载体蛋白、转运蛋白、代谢酶类以及多种受体等。这些蛋白质决定了药物的吸收、分布、靶向作用、药理作用、代谢和排泄等过程。因此,很多基因中的多种多态性都会影响药物的作用,要在全基因组寻找与药物有关的基因。

我们知道基因组内有几千种受体基因。但很多受体基因是由同一祖先基因进化而来,因此那些基因之间有较高的同源性。其结果,一种药物可能不只是与一种受体结合,而是与几种受体都有相互作用。例如氯丙嗪就与几种多巴胺能受体,肾上腺素能受体和 5-羟色胺能受体都有作用。因此多种基因的多态性都会影响药物反应。

目前寻找疾病相关的基因主要是采用连锁分析的方法,即应用呈多态性的微卫星(microsatellite)标志对家系成员进行基因分型,然后根据哪个微卫星标志与疾病相关,再去找与该微卫星连锁的基因,并最终分析确定相关基因。但由于要找到家系中每个成员服用药物反应的记录非常困难,看来连锁分析不适用于寻找药物相关基因。另一种在人群无相关个体中分析染色体某一区域与某种表型的相关研究方法则看来比较可行。但由于在开放人群中无关个体亲缘关系比家系成员要远得多。因此需要 10 万个或更多的多态性标志。这样,原来小卫星 VNTR(variable numbers of tandem repeat)和短串联重复序列(short tendem repeat,STR)的多态性位点都不够用。幸好,近几年研究的单核苷酸多态性(single nucleotide polymorphism,SNP)可以用于此目的。估计在全基因组中约有$(3\sim10)\times10^6$个 SNP。目前在基因组数据库中登录的 SNP 位点有 140 多万。大约其中 1% 有功能相关意义。那么这几万个多态位点就可作为药物基因组学研究的相关位点。

(二) 个性化用药时代

一般来说,医师根据人群中统计的平均适用药物剂量来进行治疗。这种"一种药物适合所有人"的现状终将改变,经过药物基因组研究后,医师能够选择有效的药物和患者独特的遗传学特征相匹配。这包括将具有相同表型的患者根据其不同的疾病遗传学变异、药物反应性分成更小的群体。针对这种小群体的用药比起针对广泛人群的用药可以更有效、毒性更低。

这是因为在临床上,尽管疾病的症状看起来是一样的,但个体之间的差异,特别是患者自身多基因网络作用,使得药物在每个个体身上会产生不同的疗效和毒性。药物基因组学有时可以将患者不

同的反应同他们不同的基因变异联系起来，人们希望以此推进新药的开发，并且确定某一种药物特定适应的人群，这样也可以将过去一些由于副作用而被“宣判死刑”的药物重新投入使用，不同的是，这次是使用在适合的人群中。

在过去的20多年中，人们越来越发现遗传上的杂合子同药物反应的不同有着密切的关系。有些药物在一些患者身上能够达到其他人所没有的良好效果，而在另一些患者身上会引起明显的毒性反应。有许多证据表明药物反应是基因控制的，和年龄、营养状况、健康状况和环境以及协同治疗一同起着很重要的作用。为了使个体化药物治疗达到一个好的预期转归，人们必须考虑到患者所属的地理和种族区域人群不同的药物反应模式。

药物基因组学由于在DNA分析上具备高输出技术，可以更个性化地进行有预期地用药。判明个体独特的疾病易感性和药物反应的遗传特征会对我们理解疾病的病理、选择有针对性的治疗方案产生深远的影响，也可以在为个体选择药物剂量时提供信息。这样我们就有可能真正达到理想化的个性化治疗，通俗地说就是“针对某个患者进行的适当剂量的适当药物治疗”。

（三）基因芯片（chip）及蛋白质分析技术

由于在基因组范围内进行药物相关基因的研究，工作量是巨大的。比如研究某一药物的反应需在1000个患者中进行，对10万个标志进行基因分型，就需要做1亿个基因分型，因此用常规的方法显然不可能，必须采用新的技术方法。近些年正在发展的基因芯片是解决问题的方法之一。将各种标志的探针点在膜上或玻璃片上，然后与个体标本DNA杂交，就可迅速定出某一个体的所有基因型。然后与药物反应的表型进行相关分析，就可确定哪些多态性与此相关。这样的研究甚至在不清楚药物的作用机制情况下也可进行，待找到相关基因多态性再与功能联系起来分析。国外现已有公司将6万个SNP标志放在基因芯片上进行临床药物试验的分析。将来发现更多的SNP和芯片技术的进一步提高将大大扩展这种基因组扫描方法的应用前景。

完成了人类基因组图谱和SNP图谱后，利用大量药物遗传学研究进行SNP基因型检测会是药物开发过程中不可缺少的一步。Genometrix研发出一种低成本、中密度阵列，使用了复合毛细管印刷技术（multiplex capillary printer）和高速机器人技术，每个阵列含有256个元件，96个阵列可以排列在同一片板上。在每个阵列中SNP的基因型、野生型和特定变异探针紧密排列，含有来自基因组DNA序列的目的多态型的靶序列经过复合PCR反应，再分析杂交后的荧光信号，用一个阵列就可以分析多个SNP，见图10-6。

应用基因芯片技术也可检测某种细胞或组织的全部基因表达图谱。比较不同患者服药后基因表达图谱的差异，就可分析药物在不同个体中作用机制及代谢等方面的差异。组织基因表达图谱分析特别适用于癌症患者。可将组织活检标本或手术切除标本进行mRNA抽提和分析。肿瘤中改变的基因表达可指导有效药物的选择，并避免不必要的毒性和无效药物的应用。如乳腺癌患者中癌基因*erb-B2*扩增，就可预测用环磷酰胺-氨甲蝶呤-5-氟尿嘧啶（CMF）佐剂治疗效果较好。在其他的疾病中，转录表达图谱也可用于确定一些候选基因的改变。如从气管中取得的上皮和炎症细胞中转录图谱可确定哮喘治疗的药物反应等等。在进行临床检测转录图谱中较困难的是取得单一种类的细胞。许多标本往往掺杂着各种细胞，对转录图谱有很大干扰。现已有一种激光捕获微量切片（laser capture microssection，LCM）方法，可从组织切片中只取单个细胞，然后进行mRNA扩增和分析。这就保证了转录图谱分析的准确性。同时也只需极少的临床样本，有利于检测的开展。

现在已有报道通过基因芯片技术快速检测G-6-PD基因突变，进行G-6-PD缺乏症的筛选，应用于临床的预防和治疗。

基因的最终产物是蛋白质，因此蛋白质也是药物基因组学研究的对象。现已在基因组范围开展的蛋白质分析研究技术是二维凝胶电泳结合质谱分析技术。将细胞所有蛋白质抽提出来，在聚丙烯酰胺凝胶上进行PAGE和等电聚焦双向电泳，然后应用高灵敏度的质谱仪分析蛋白质斑点。

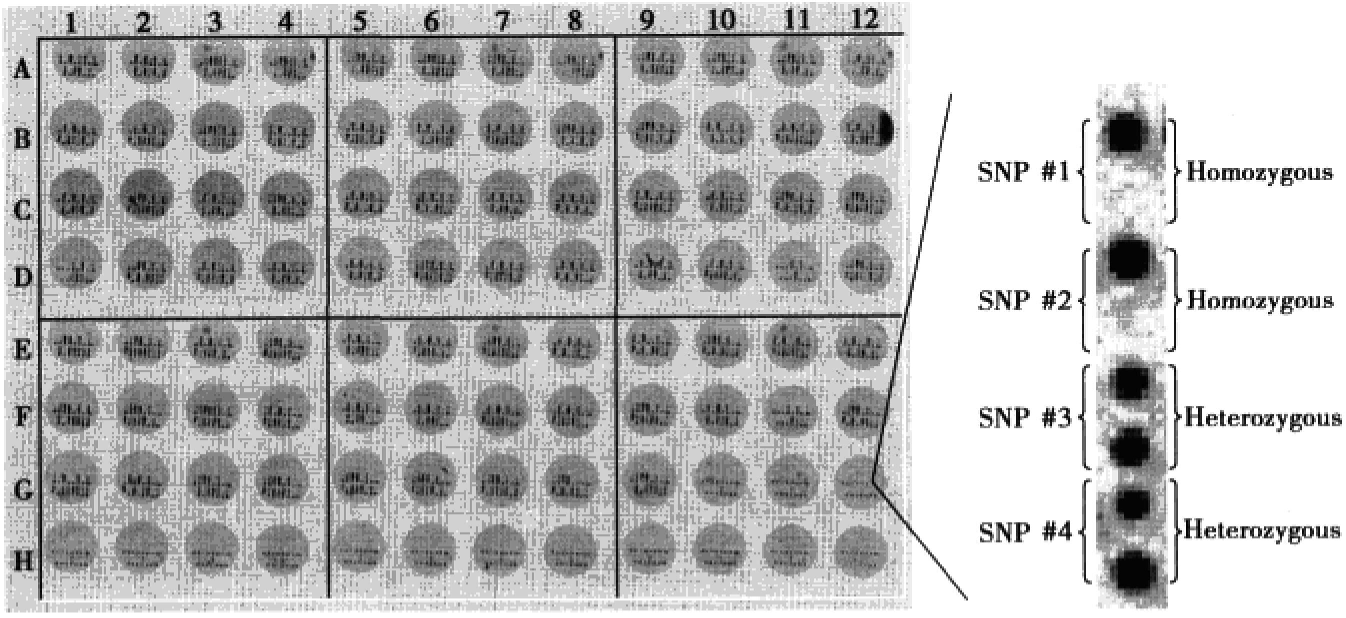


图 10-6　使用 Genometrix's VistaArray™96 孔阵列进行的微阵列基因型测定

在平板玻璃表面每个阵列含有最多 256(16×16)元件(features)。每个 SNP,野生型以及其他不同的探针被设计粘附在阵列上。用荧光扫描仪分析杂交信号。与一个等位基因探针杂交的是纯合子基因型,与野生型和变异型探针都杂交的是杂合子基因型

比较用药前后细胞表达蛋白的差异,也可了解药物作用效果等情况。而不同个体中表达图谱的差异也可作为用药的依据。但这种方法自身存在不少问题。除了仪器制造复杂之外,它们不能很好分离强电荷或低质量蛋白质,还不能分离具有疏水区域(例如跨越细胞膜的区域)的蛋白质。这是一个重大的缺陷,因为跨膜受体是重要的药物标靶。因此以后临床大量应用可能性不大。但蛋白质芯片技术则可能较适合于临床应用。它是将抗体固定在膜片或玻片上,然后与细胞表达的各种蛋白质分别结合。一些 SNP 位点可能正好在蛋白质抗原决定簇,因此不同的基因型可产生不同的抗体。检测哪种抗体与蛋白质结合,就可了解个体是哪种基因型,同时也就了解了蛋白质表达图谱。

从理论上讲,了解人类基因组中结构基因所表达的每一种蛋白质的确切结构形成将有助于药物设计者设计出化合物来填充蛋白质上的位置,从而要么刺激要么阻止蛋白质相互作用。这种工作就是通常所说的合理药物设计,但至今尚未取得广泛的成功,因为目前只有 1% 的人类蛋白质的结构已被确定。不过一旦人类蛋白质的结构全部搞清以后,风行一时的将是蛋白质而不是基因。

(四) 研究前景

药物基因组学在经过大规模研究后,我们能从中得到多少收益呢？首先,我们可以通过基因分型而避免极少数患者因个体差异而受到严重的药物毒性反应,甚至丧失生命。药物也可预先进行设计,不要造成太大的个体差异。如设计的药物在代谢中不需与细胞色素氧化酶 P450 中 CYP2D6 相互作用,以防止在 PM 患者中引起毒性反应。

有些药物常常只在少数患者中有很好的治疗效果,在大部分人中治疗效果很差,甚至没有效果。如晚期大肠癌手术加化疗治愈率仅 20% ~40%,大部分患者对化疗反应不良。因此有必要在患者治疗前为其选择最合适的药物。可能建立 SNP 基因芯片在治疗前分析患者基因型可达到此目的。关键是基因芯片必须含有经过基因组扫描获得关键信息的 SNP 位点。这样就可判断哪些患者对哪种药敏感,或对哪种药耐受等。选择最合适的药物治疗就会事半功倍。

但是,有些药物的作用可能受到很多基因的影响,每个基因又有很多多态性部位,要确定哪些多态性部位决定临床药物有效性会有些困难。如抗癌药物 5-氟尿嘧啶在治疗肠癌时可被近 40 种不同的酶激活或灭活。要确定哪个酶,哪几种多态性部位是关键还需做大量、细致的研究工作。

药物基因组学也可在药物开发研究方面节省很多费用,因为原来有些药物疗效很好,只是因为在一些人中有严重的副作用而不得不放弃。采用药物基因组学技术找出是哪种基因型个体有

副作用，就可在其他个体中应用这种药物。这样就大大增加了研究药物的可用性而减少了研究开发费用。

药物基因组学在接下来的研究中还有大量工作要做。当务之急是要系统地鉴定一整套频率大于1%的遗传多态性标记。在此基础上建立与药物反应差异相关的数据库。这套数据库既可对患者服药反应的表型进行相关基因型评估，又可为研究其他各种表型作进一步的研究，从而不断充实数据库。目前重要的工作是要鼓励个人和家庭建立各种细胞、组织和体液等样品的数据库，以便对疾病和药物作用进行精细的研究并指导以后的治疗等。

几乎在人类基因组计划研究的同时，许多国家的学者正在对其他生物的基因组进行分析研究，试图在人与一些简单生物的比较中发现和寻找能用于人类的新药。在许多情况下，人类同果蝇、线虫、酵母和小鼠等模式生物的基因所编码的蛋白质十分相似。任何一个人类基因在线虫或果蝇的基因组中找到足够相似的对应者的概率约为50% ~90%，这样我们可把这些简单生物作为模式来进行研究，寻找适用于人类的药物。目前对果蝇的基因组的测序已全部完成。在268种已知的人类疾病基因中，60%在果蝇中都有对应者，全部果蝇蛋白质的50%（约7000种）与已知的哺乳动物的蛋白质相似。与人类对应的果蝇基因之一是P53，这种类似性使果蝇成了研究人类癌肿遗传的"良好对等物"，未来的抗人类肿瘤的新药很可能先在果蝇身上进行的研究中取得突破。秀丽线虫的基因组测序也已完成，结果发现，其1/3蛋白质（约6000种以上）与已知的哺乳动物的蛋白质相似。目前一些公司正利用这些长约1mm的线虫进行自动筛选试验以寻找新的治疗糖尿病的药物，他们把因胰岛素受体基因突变而生长受阻的线虫放到药丸大小的凹陷里，然后加入不同的化学品到实验的凹陷里，以检测哪些能恢复线虫的生长，从而表明这些化合物能通过有缺陷的受体。由于许多糖尿病患者的细胞不再响应胰岛素，于是这些化合物可以作为治疗糖尿病新药的基础。酵母是第一种被测序的有核生物。约38%的酵母蛋白质（2300种左右）与已知的哺乳动物蛋白质相似，这使酵母成为一种非常好的癌症研究生物模型。美国的科学家已经从酵母的研究中了解了许多有关细胞分裂和DNA修复的大量资料，这些资料对肿瘤治疗的研究是重要的，他们利用酵母来解释一些现用的抗肿瘤药物是如何作用的，已有的发现之一就是普通的化学药物cisplatin对杀死癌细胞特别有效，而这些细胞在修复DNA的能力方面存在特定的缺陷。由于所有的新药最终都必须在哺乳动物身上进行试验，而小鼠往往是首选的对象。因为从基因组的角度来说，小鼠与人是很接近的。而从已经识别出的小鼠蛋白质来看，与已知的人类蛋白质的结构很相似。作为一种理想的模式生物，小鼠的基因组测序工作正在加紧进行，这将大大推进药物基因组学的研究。

总之，药物基因组学在药物设计、制造和应用方面正酝酿着一场根本性的革命。将目前依据患者群共性的药物治疗转向今后根据不同人群及不同个体遗传特征来设计和制造药物，从而最终达到个体化治疗（individualized therapy）的水平，为人类认识自我，保持健康，延长寿命作出贡献。

Summary

Pharmacogenetics is the science and study of drugs and their administration in relationship to genetics Although variability in drug response may depend on many factors, genetic factors have been shown to play a significant role. Genes that influence drug response code for drug receptors or targets for examples, or enzymes that are involved in drug metabolism or drug clearance. Pharmacogenetics deals with genetic predisposition to drug reactions, effectiveness of drug treatments on individuals depending on genetic predisposition, and with prediction of disease susceptibility based on genetic profile.It is the study of how our unique genetic make-up makes every individual respond differently to medicines. It helps physicians assess which drugs are suitable for each person rather than using a 'one size fits all' approach. Pharmacogenetics enables us to re-

duce unnecessary side affects from medicines and ensure effective and safe treatment. The field is still in its infancy, so pharmacogenetics is unlikely to impact much on clinical practice in the near future. But in the longer term it could improve the safety and efficacy of medicines. With advancements in scientific technology, Pharmacogenetics is becoming increasingly important both in administration of medication and the treatment of disease

（陆振虞）

参考文献

1. 贺林，李胜. 基因组与环境以及药物开发的联系//贺林. 解码生命. 北京：科学出版社，2000，277-285
2. Meyer UA, Zanger UM. Molecular mechanism of genetic polymorphisms of drug metabolism Annu. Rev Pharmacol Toxicol, 1997, 37: 269-296
3. Grant DM, Morike K, Eichelbaum M, et al. The slow acetylator phenotype is caused by decreased or absent arylamine N-acetyltransferase in human liver. J Clin Invest, 1990, 85: 968-972
4. Gaedigk A, Blum M, Gaedigk R, et al. Deletion of the entire cytochrome P450 CYP2D6 gene as a cause of impaired drug metabolism in poor metabolizers of the debrisoquine/sparteine polymorphism. Am J Hum Genet, 1991, 48: 943-950
5. Kleyn PW, Vesell ES. Pharmacogenomics: Genetic variation as a guide to drug development. Science, 1998, 281 (5384): 1820
6. Evans WE, Relling MV. Pharmacogenomics: Translating functional genomics into rational therapeutics. Science, 1999, 286: 487-491
7. Gonzalez FJ. Characterization of the common genetic defect in humans deficient in debrisoquine metabolism. Nature, 1998, 331: 442-446
8. Sadee W. Genes and the pharmaceutical sciences. Parmaceut Res, 1988, 5: 199-200
9. Evans W E. Pharmacogenomics: marshalling the human genome to individualise drug therapy. Gut, 2003, 52(2): 10-18
10. Olivier Tribut. Pharmacogenomics. Med Sci Monit, 2002, 8(7): 152-163
11. Laviero Mancinelli. Pharmacogenomics: The promise of personalized medicine. AAPS Pharma Sci, 2000, 2(1): 1-13
12. Michael MS. Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. Clinical Chemistry, 2001, 47(2): 164-172
13. Emilien G.. Impact of genomics on drug discovery and clinical medicine. Q J Med, 2000, 93: 391-423

第十一章　肿瘤遗传学

肿瘤（tumor）泛指由一群生长失去正常调控的细胞形成的新生物（neoplasm），分为良性肿瘤（benign tumor）和恶性肿瘤（malignant tumor），上皮来源的恶性肿瘤又称为癌症（cancer），间叶组织来源的称为肉瘤（sarcoma）。通常，肿瘤细胞是一个积累了不同基因突变的体细胞，这些突变共同作用导致了细胞增殖的失控，结果形成大量细胞的集合体——肿瘤。肿瘤形成后可在原位继续生长，也可转移并进入其他组织器官，而侵袭到其他部位的肿瘤恶性程度较高。肿瘤细胞持续生长将出现严重的组织损伤和器官衰竭，最终导致死亡。目前已发现的恶性肿瘤几乎涉及了所有类型的细胞、组织及器官系统，其中约 85% 是癌，起源于上皮细胞；2% 是肉瘤，来源于间叶细胞；约 5% 为淋巴瘤，来源于免疫系统，特别是脾及淋巴结的白细胞；约 3% 为白血病，来源于骨髓造血细胞。

肿瘤是一种体细胞遗传病，是相关基因结构与功能改变的结果。这些基因参与控制细胞生长

和凋亡过程、细胞增殖或损伤修复。致癌突变主要分两种：一种是种系突变，直接由双亲传给子女，出现在后代所有体细胞和生殖细胞中并向下传递，有5%～10%的肿瘤由种系突变引起，称为遗传型（发生在家族中）肿瘤；另一种称为散发突变，可发生在人一生中的任何时候，而不是从双亲继承而来，引起这些突变的原因可能是烟草、紫外线照射、病毒、年龄及其他因素，散发突变引起的肿瘤称为散发型肿瘤，相比遗传型肿瘤常见得多。在多数遗传型及所有散发型恶性肿瘤中，肿瘤的始动和发展需要多次体细胞突变。肿瘤的遗传基础十分复杂，因为突变基因的不同组合可发生相同类型的肿瘤。一些基因的突变可见于多种类型的肿瘤，而另外一些基因突变只限于少数特定类型的肿瘤。

进入21世纪，人类对肿瘤病因的了解依然有限。并且，由于感染性疾病得到有效控制及人类寿命的延长，患恶性肿瘤的风险与几率大大增加。此外，环境污染、压力增加及不良生活方式的影响（如吸烟等），也在一定程度上增加了肿瘤的发生率。然而，多数恶性肿瘤的发生并没有明显的决定性因素，其中遗传基础起到了重要的作用。如果仅仅是一种单基因病，那么理解恶性肿瘤是如何始动并设计治疗方案会更加容易。但事实并非如此简单，当诱导细胞分裂的过程出现失控、细胞分裂周期的调控缺陷和/或程序性细胞死亡（凋亡）的控制被破坏时均会导致肿瘤的发生。

第一节　染色体异常与肿瘤

自从细胞遗传学技术开始应用于人类恶性肿瘤的研究以来，其中频发且复杂的染色体畸变（chromosomal aberration）一直备受关注。染色体研究证实了肿瘤是一种体细胞遗传病，并发现在多数人类恶性肿瘤中均伴有染色体数目或结构的异常。即使在分子生物学技术于肿瘤研究中广泛应用并取得突破性结果的今天，染色体改变与恶性肿瘤发生发展之间关系的探索仍然是人类恶性肿瘤研究中的一个重要领域。

一、肿瘤细胞的克隆演进与染色体变异

多数肿瘤细胞具有染色体异常（图11-1）。在一个肿瘤的细胞群体中，通常是由单克隆构成，即单克隆起源，但肿瘤生长演进过程中会出现异质性（heterogeneity），演变为多克隆性，其中占主导数目的克隆构成肿瘤干系（stem line），干系肿瘤细胞的染色体数目称为众数（modal number）。多克隆细胞群肿瘤中占非主导数目的克隆称为旁系（side line）。由于细胞内外条件变化，干系与旁系地位可以相互转变，甚至单克隆细胞群亦可以发展为多克隆肿瘤细胞群，这种现象称为肿瘤细胞的克隆演进。

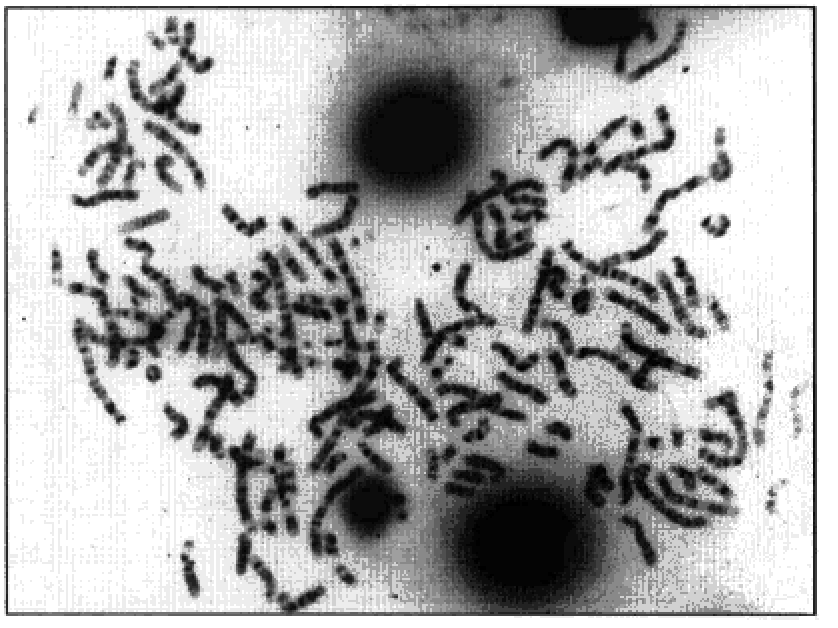


图11-1　人类恶性肿瘤细胞中染色体数目与结构异常

肿瘤细胞的核型多伴有染色体数目的改变，大多是非整倍体（aneuploid），其中包括超二倍体（hyperdiploid），即比46多一条或几条染色体；亚二倍体（hypodiploid），即比46少一条或几条染色体；亚三倍体（hypotriploid），即比69少一条或几条染色体；亚四倍体（hypotetraploid），即比92少一条或几条染色体。染色体数目在三倍体以上者称为多倍体（polyploid）。实体瘤染色体数目多为三倍体左右。此外，肿瘤细胞核型中亦频发染色体的结构异常。在肿瘤的发生发展过程中，由于肿瘤细胞的增殖失控等原因，导致细胞有丝分裂异常并产生部分染色体断裂与重

排,形成了一些结构特殊的染色体,称为标记染色体(marker chromosome)。标记染色体的形成可能是随机的,但有一小部分能够在肿瘤细胞中稳定遗传,称为特异性标记染色体,这是一种非随机事件。这些特异性标记染色体与肿瘤的恶性程度及转移能力密切相关。

二、Ph 染色体的发现及其意义

1960 年,Nowell 和 Hungerford 在慢性粒细胞性白血病(chronic myelocytic leukemia,CML)中发现了一条比 G 组染色体还小的异常染色体,因在美国费城(Philadelphia)发现而被命名为 Ph 染色体(Philadelphia chromosome)。经显带技术证明 Ph 染色体是 t(9;22)(q34;q11.2)(图 11-2)。约 95% 的慢性粒细胞性白血病细胞携有 Ph 染色体,它可以作为 CML 的诊断依据。

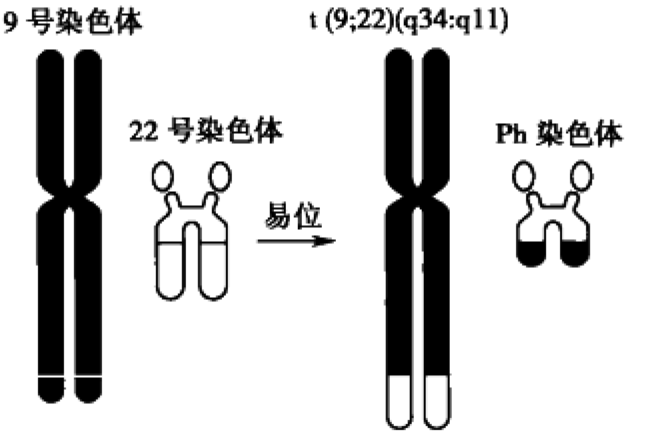


图 11-2 Ph 染色体构成示意图

Ph 染色体的发现首次证明了一种染色体畸变与一种特异性肿瘤之间的恒定关系,故被认为是肿瘤细胞遗传学研究的里程碑。

三、肿瘤中其他特异性标记染色体

除 Ph 染色体外,还发现了一些非随机存在的特异性标记染色体,它们的特点是与特定的肿瘤相关,但检出率并不比 Ph 染色体高,其中包括 Burkitt 淋巴瘤中的 t(8; 14),脑膜瘤中的 $22q^-$ 或 −22,视网膜母细胞瘤中的 $13q14^-$,神经胶质细胞瘤中的 1 号和 22 号染色体变异,慢性髓细胞性白血病急性变呈现的 +8 和 $17q^+$,急性白血病中的 −7 或 +9,Wilms 瘤中的 11 号染色体短臂中间缺失 del (11p13→p14),结肠息肉中的 +8 或 +14,小细胞肺癌的 3 号染色体短臂中间缺失 del (3p14→p23),肺腺癌与肺鳞癌中 6 号染色体长臂末端缺失 del (6q23→qter)(图 11-3),胃癌中的 22 号、17 号和 18 号染色体畸变,黑色素瘤中的 del (1)(p12q12) 改变和 1 号与 6 号染色体之间的重排,鼻咽癌的 t (1; 3)(q41; p11) 以及乳腺癌中的涉及 1 号染色体长臂的易位等等。

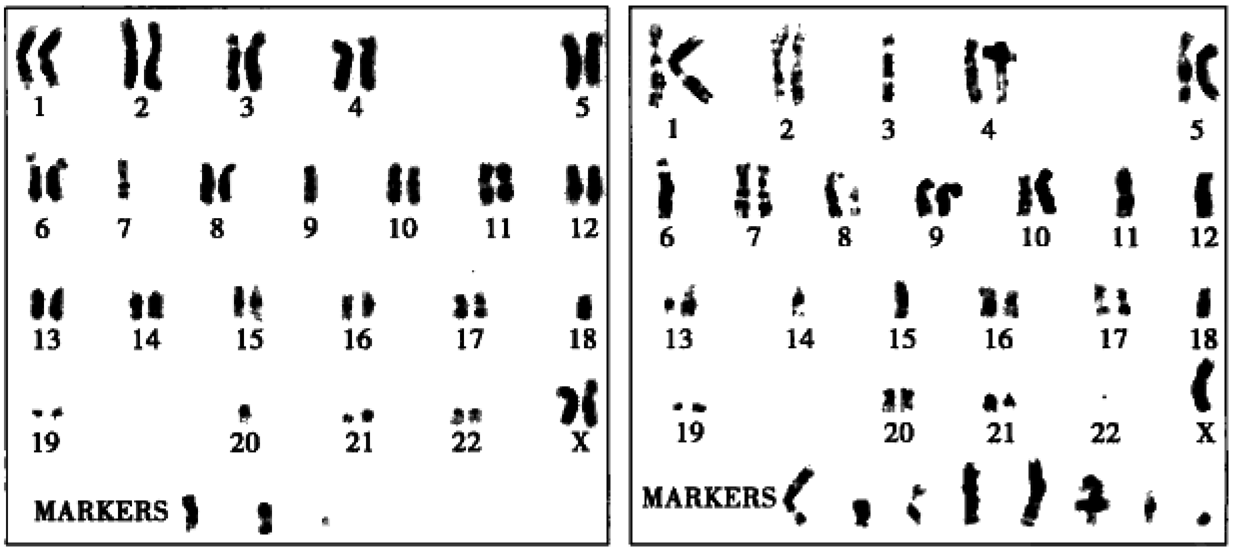


图 11-3 原发性非小细胞肺癌细胞中的染色体改变

第二节 癌 基 因

20 世纪初,Boveri 提出了"在分子水平上,肿瘤是由于细胞 DNA 损伤引起的"这一观点。随后研究发现,在反转录病毒的基因组中除了编码病毒本身复制所必需的病毒核心蛋白、外壳糖蛋白及反转

录酶等基因外，还包括一个能引起动物宿主细胞恶性转化的基因。这种基因就是癌基因（oncogene），又名转化基因，是包括人类在内的动物细胞及致癌病毒固有的一类基因，它们一旦异常活化便能促使人或动物的正常细胞发生癌变。

一、癌基因的发现及识别

癌基因的发现可追溯到20世纪初对动物致癌病毒的研究。人们注意到患白血病家禽的细胞提取物注入正常家禽体内可引起白血病，并发现肿瘤细胞中含有病毒。

1910年Rous发现一种病毒可使禽类产生肿瘤。此病毒能使鸡胚成纤维细胞在培养中转化，再给鸡接种后还能诱发肉瘤，称为Rous肉瘤病毒（Rous sarcoma virus，RSV）。后经证实RSV是一种RNA反转录病毒，除含有病毒复制所需基因外，还含有一种特殊的转化基因，不仅能使培养的细胞转化并呈现恶性表型，也能在动物中引发恶性肿瘤，这种基因被称为病毒癌基因（viral oncogene，v-oncogene）。这项研究的深远意义使Rous于1966年获诺贝尔生理学或医学奖。

1969年Huebner和Hodaro提出癌基因假说（oncogene hypothesis），认为所有的细胞中都含有致癌病毒的全部遗传信息，这些遗传信息代代相传，其中与致癌有关的信息称为癌基因。通常情况下癌基因处于被阻遏状态，只有当细胞内的有关调节机制遭破坏时癌基因才表达，进而导致细胞癌变。

1970年Temin等发现致癌的RNA病毒中存在着一种反转录酶，于是提出了原病毒假设（provirus hypothesis），认为RNA病毒通过反向和正向转录以及与宿主细胞DNA发生交换或重组，形成癌基因。Temin于1975年凭此研究获得诺贝尔生理学或医学奖。

自20世纪70年代起，癌基因研究得到快速发展，并取得了一些实质性突破。1970年Martin用RSV的温度敏感突变体证实RSV基因组内确实存在能使体外培养的正常细胞转化为癌细胞的基因，即被称为*src*的病毒癌基因，是病毒感染的正常细胞转变为癌细胞的基因。*src*是人类发现的第一个病毒癌基因。

1971年Duesberg等比较野生型RSV和突变型RSV（*src*缺失）的基因组，发现*src*基因位于野生型RSV基因组RNA的3′端，突变型RSV除了无致癌作用外一切正常，提示*src*对RSV的生长和增殖并非必需。在观察了RSV及其亲缘病毒后，大多数遗传学家认为，只有RSV才是真正的自然病毒，而它们的亲缘病毒则由于某种缘故失去了*src*基因，成为相关致癌能力有缺陷的突变病毒。那么RSV中*src*基因从何而来？

Evarmus和Bishop合作研究很快证实了*src*基因的真正来源：1975年他们用发明的检测病毒及细胞基因组中*src*基因的新技术（用*v-src*探针作Southern杂交）分析正常鸡细胞及感染RSV的鸡细胞基因，发现在未感染病毒的细胞和感染病毒的细胞中都有与*v-src*非常相似的*src*基因存在，说明正常鸡细胞在感染RSV之前，就已拥有至少一个*src*基因，*src*基因本来就是一个正常的鸡细胞基因，当RSV病毒在受感染的鸡细胞内生长时，在某种条件下，通过某种遗传事件，把细胞的一个*src*基因并入自己的病毒基因组，使一个正常的细胞基因被再造成强有力的致癌基因。这是第一个被确认的“细胞癌基因”（cellular oncogene，c-oncogene）。不久，在鸟类和包括人类在内的脊椎动物基因组内也相继发现了*src*基因，表明*src*基因是所有脊椎动物正常遗传“设备”的一部分。至此，Huebner和Hodaro有关癌基因的假设在许多实验中得到了证实，病毒癌基因是通过转导作用从宿主细胞原癌基因中掳获的，即癌基因起源于动物。Evarmus和Bishop小组把正常细胞中的*src*基因称作原癌基因（pro-oncogene，POG），意指其具有在适当环境下被激活变为癌基因的潜力。在包括人类在内的所有脊椎动物基因组中，均能发现各种类型的原癌基因，每种原癌基因都和*src*一样，广泛分布在动物王国内部。癌基因是原癌基因的变异形式。Bishop由于在癌基因研究中的贡献而获1989年诺贝尔生理学或医学奖。

二、癌基因、原癌基因及其功能

原癌基因多编码调控细胞生长的蛋白质，其通过异常激活转变为癌基因并出现功能或表达的改变，诱导易感细胞形成肿瘤。根据原癌基因蛋白产物的功能及生化特性，可将癌基因分为五类：生长因子、生长因子受体、信号转导因子、转录因子及其他，如程序性细胞死亡调节因子。表 11-1 列出了部分根据其功能进行分类的原癌基因。

（一）生长因子

生长因子是分泌性多肽，作为细胞外信号可以刺激靶细胞的增殖。几乎所有的靶细胞都具有和相应生长因子相结合的受体。生长因子与反转录病毒癌基因之间的关系是在研究猴肉瘤病毒的 *pdgfb* 基因的过程中发现的，该病毒是从猴纤维肉瘤中提取分离出来的反转录病毒。序列分析证实 *pdgfb* 基因编码血小板源生长因子 β 链（platelet derivative growth factor B，PDGFB），这表明异常表达的生长因子具有癌蛋白的作用。

（二）生长因子受体

一些癌基因来自编码具有内源性酪氨酸激酶活性生长因子受体基因的变异。这些作为生长因子受体的酪氨酸激酶均有一个典型的蛋白质结构，包括 3 个基本区域：细胞外配体结合区域、转膜区域及细胞内酪氨酸激酶催化区域。生长因子受体是一个将信息进行单向跨膜传递的“分子工具”。生长因子与受体的胞外配体结合区相结合，使细胞内酪氨酸激酶催化区域激活，触发一系列生化反应，最终导致细胞分裂。

既然生长因子受体既在正常细胞生长过程中起着重要调节作用，也是原癌基因家族重要的组成部分，例如 *EGFR*、*ERBB2*、*CSF1R*、*KIT*、*MET*、*RET*、*ROS1* 及 *NTRK1* 等，生长因子受体基因的突变或异常表达都可以使它们转化为癌基因。

（三）信号转导因子

促有丝分裂信号由位于细胞表面的生长因子受体传递到胞核中需要经过一系列复杂的反应途径，即信号转导的级联反应。信息的传递一部分是依靠胞液中相互作用的蛋白质的逐级磷酸化，也有一部分与鸟氨酸结合蛋白及第二信使（如腺苷酸环化酶系统）相关。人类发现的第一个反转录病毒癌基因 *src* 就是信号转导因子。

许多原癌基因都是信号转导通路的组成部分，信号转导因子由于突变可转变为癌基因，使其活性不受控制，继而使细胞出现无限增殖。

（四）转录因子

转录因子是一种能够调节目的基因或基因家族表达的核蛋白。转录调控通常是由蛋白质与特异 DNA 序列或 DNA 结构域相结合介导的。转录因子多属于多基因家族，有共同的 DNA 结合区域，例如锌指结构。转录因子是信号转导途径的最后一个环节，它将细胞外信号转换为调节基因表达的效应。

在研究同源反转录病毒时发现许多原癌基因属于转录因子家族，例如 *THRA*、*ETS1*、*FOS*、*JUN*、*MYB* 及 *MYC*，其中 *FOS*、*JUN* 构成 AP-1 转录因子，AP-1 能促进很多目的基因的表达而引起细胞分裂。在造血系统肿瘤和实体瘤中，有转录因子功能的原癌基因通常是由于染色体易位而被激活。

（五）程序性细胞死亡调节因子

正常组织在细胞增殖与死亡之间的调节有一种平衡。在正常胚胎形成及器官发育过程中，程序性细胞死亡是一个重要的调节机制。成熟的组织中程序性细胞死亡的主要表现形式又称为细胞凋亡，其在形态学上有以下特性：细胞膜空泡变性、体积凝缩、胞核凝集、基因组 DNA 由于内源性核酸酶的作用而被分割为以核小体单位为大小的碎片。研究发现不受程序性细胞死亡调节的细胞可出现无限增殖并容易形成肿瘤。目前发现可以调节程序性细胞死亡的原癌基因是 *BCL2*、*MDM2*。

表 11-1 原癌基因分类

分类	原癌基因	染色体	肿瘤类型	蛋白功能
生长因子	*PDGFB*	22q12.3-q13.1	胶质瘤/纤维肉瘤	血小板衍生生长因子 B 链
	FGF3	11q13	乳腺癌	成纤维细胞生长因子家族成员
	KS3	11q13.3	Kaposi's 癌	成纤维细胞生长因子家族成员
	FGF4	11q13	胃癌	成纤维细胞生长因子家族成员
生长因子受体	**酪氨酸激酶:组成性膜蛋白**			
	EGFR	7p12.3-p12.1	鳞状细胞癌	表皮生长因子受体
	CSF1R	5q33.2-q33.3	肉瘤	集落刺激因子 1 受体
	KIT	4q12	肉瘤	干细胞因子受体
	ROS1	6q22	肉瘤	?
	MET	7q31	MNNG 处理的人软骨细胞系	造血生长因子受体
	NTRK1	1q21-q22	结肠/甲状腺癌	神经生长因子
	ERBB2	17q21.1	成神经细胞瘤/乳腺癌	?
	RET	10q11.2	甲状腺癌	GDNF/NTT/ART/PSP 受体
	缺乏蛋白激酶活性的受体			
	MAS1	6q24-q27	表皮样瘤	血管紧张素受体
信号传递因子	**细胞质酪氨酸激酶**			
	SRC	20q12-q13	结肠癌	蛋白酪氨酸激酶
	YES1	18p11.3	肉瘤	蛋白酪氨酸激酶
	FGR	1p36.2-p36.1	肉瘤	蛋白酪氨酸激酶
	FES	15q26.1	肉瘤	蛋白酪氨酸激酶
	ABL1	9q34.1	慢粒	蛋白酪氨酸激酶
	膜相关的 G 蛋白			
	HRAS	11p15.5	结肠癌、肺癌、胰腺癌	GTP 酶
	KRAS	12p12.1	AML、甲状腺癌、黑色素瘤	GTP 酶
	NRAS	1p13.2	肉瘤、黑色素瘤	GTP 酶
	GSP	20	甲状腺腺癌	Gs alpha
	GIP	17q21.3-q22	卵巢癌、肾上腺癌	Gi alpha
	GTP 酶交换因子(GEF)			
	MCF2	Xq27	弥漫性 B 细胞淋巴瘤	Rho 和 Cdc42Hs 的 GEF
	VAV1	19p13.3-p13.2	造血细胞	Ras 的 GEF
	细胞质丝氨酸/苏氨酸激酶			
	MOS	8q11	肉瘤	蛋白激酶(丝氨酸/苏氨酸)
	RAF1	3p25	肉瘤	蛋白激酶(丝氨酸/苏氨酸)
	PIM1	6p21.2	T-细胞淋巴瘤	蛋白激酶(丝氨酸/苏氨酸)
	细胞质调控因子			
	CRK	17p13.3		SH-2/SH-3 接头

续表

分类	原癌基因	染色体	肿瘤类型	蛋白功能
转录因子	*MYC*	8q24. 12-q24. 13	肉瘤/髓细胞瘤	转录因子
	MYCN	2p24. 1	神经母细胞瘤、肺癌	转录因子
	MYCL1	1p34. 3	肺癌	转录因子
	MYB	6q22	成髓细胞性白血病	转录因子
	FOS	14q24. 3	骨肉瘤	转录因子 API
	JUN	1p32-p31	肉瘤	转录因子 API
	SKI	1p36. 3	癌	转录因子
	REL	2p13-p12	淋巴性白血病	突变型 NFKB
	ETS1	11q23. 3	成红细胞增多病	转录因子
	ETS2	21q22. 3	成红细胞增多病	转录因子
	THRA	17q11. 2	成红细胞增多病	T3 转录因子
	THRB	3p24. 3	成红细胞增多病	T3 转录因子
细胞凋亡调节因子	*BCL2*	18q21. 3	B 细胞淋巴瘤	抗凋亡蛋白
	MDM2	12q14. 3-q15	肉瘤	P53 调控蛋白

三、癌基因的激活机制

癌基因的激活与细胞原癌基因的遗传特性改变有关。这些遗传特性变异的结果使细胞获得了一定的生长优势。在人类肿瘤中癌基因激活有 3 种机制：突变、基因扩增（gene amplification）及染色体重排（chromosome rearrangement），这些机制或改变原癌基因的结构，或增加其表达量。

（一）突变

突变的原癌基因通过其编码的蛋白质结构的改变而激活。这些变异通常涉及一些关键的蛋白调节区域，导致突变蛋白不受调控并出现持续性激活。各种类型的基因突变如碱基替换、缺失或插入，都有可能激活原癌基因。例如，反转录病毒癌基因，经常由于缺失而被激活。此外还有 *EGFR*、*KIT*、*ROS1*、*MET* 及 *TRK* 癌蛋白由于氨基末端配体结合区的缺失而被激活。但在人类肿瘤中，典型的癌基因突变多数是由于碱基替换（点突变）导致的，即编码蛋白中仅有一个氨基酸的变异。

人类肿瘤的早期研究发现：在原癌基因 *RAS* 家族（*KRAS*、*HRAS*、*NRAS*）中，经常可以检测到点突变。据统计，在随机挑选的肿瘤中大约每 15% ~20% 的病例存在一个 *RAS* 基因突变。*KRAS* 的突变在恶性肿瘤中尤为常见。研究表明，在 30% 的肺腺癌、50% 的结肠癌及 90% 的胰腺癌中存在 *KRAS* 的突变。*NRAS* 的突变主要发生在造血系统的恶性肿瘤中，在急性粒细胞白血病及骨髓异常增生综合征中 *NRAS* 的突变率为 25%。大部分甲状腺癌中同时存在 3 种癌基因（*KRAS*、*HRAS*、*NRAS*）的突变，而不是单一 *RAS* 基因的突变。特别是在已分化的滤泡型甲状腺癌中，三种 *RAS* 基因突变的联合作用尤为明显。*RAS* 基因突变的结果可使具有信号转导功能的 RAS 蛋白被激活。

（二）基因扩增

大规模的突变涉及染色体的部分重复或缺失。一旦细胞获得染色体某些位点（通常含一个或多个癌基因以及毗邻的遗传单位）的多个拷贝（常常 20 个以上），就出现了基因扩增，即基因组中某个基因拷贝数的显著增加。基因通过其在基因组内异常扩增，引起核型改变，包括均质染色区（homogeneous staining regions，HSRs）（图 11-4）和双微体（double minute chromosomes，DMs）（图 11-5）等。HSRs 是缺少正常深、浅染色区的染色体片段，DMs 是典型的无着丝粒的微小环状遗传结构。DMs、HSRs 代表基因组 DNA 高度扩增区，其中含有至少几百个相同拷贝。扩增使基因的表达增高，进而为细胞生长提供了优势。

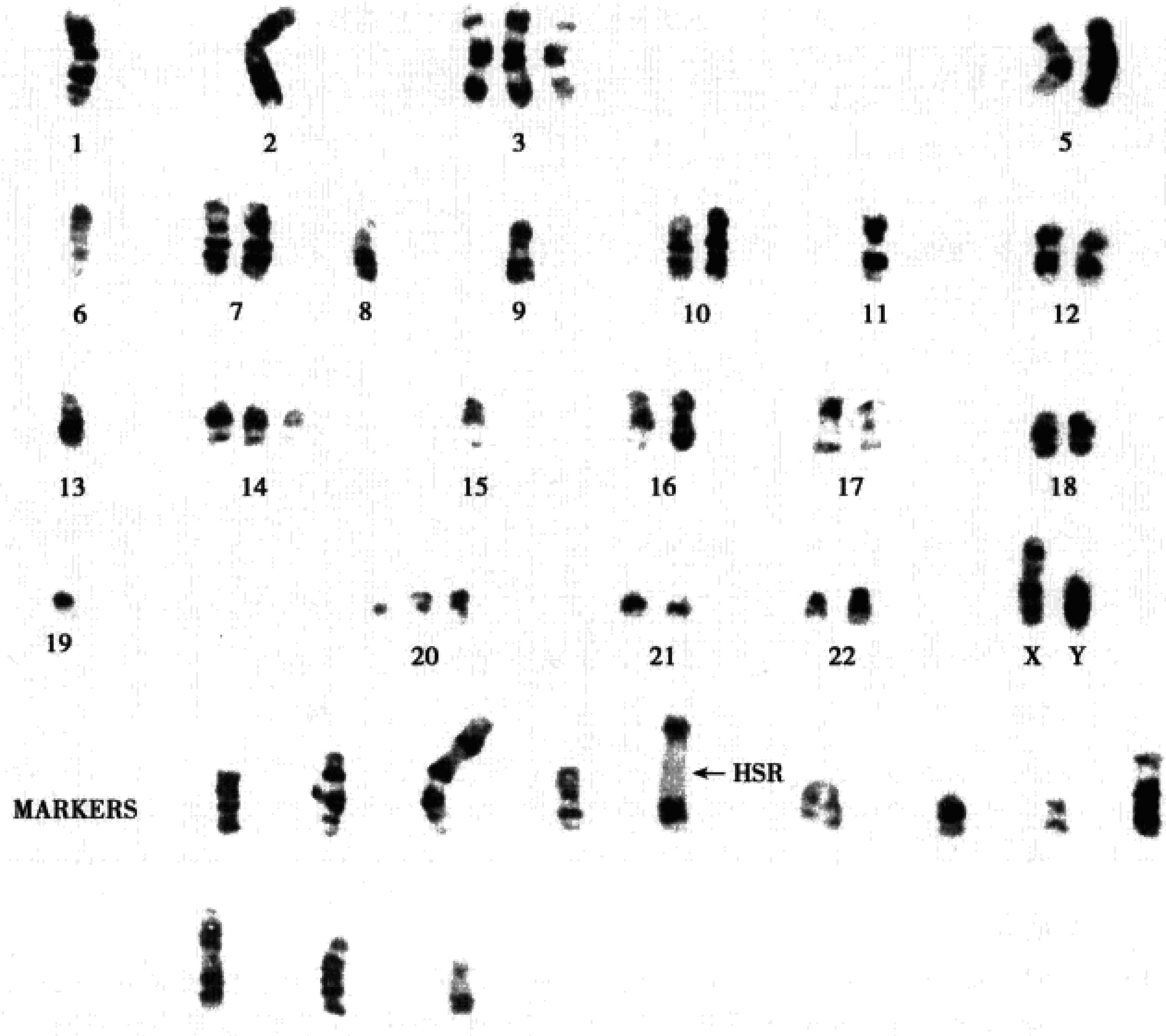


图 11-4　人类恶性肿瘤细胞染色体上的均质染色区（箭头示）

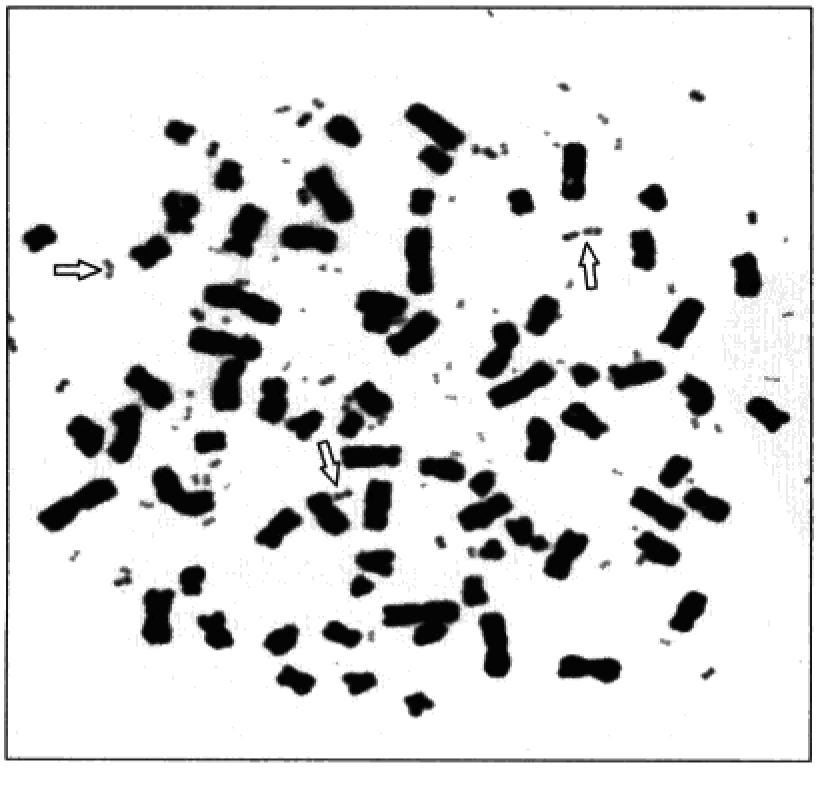


图 11-5　人类恶性肿瘤细胞核型中的双微体（箭头示）

人类肿瘤核型中频繁出现 HSRs 与 DMs，表明在肿瘤中某些原癌基因的扩增是很常见的，其中 3 个原癌基因家族：*MYC*、*EGFR* 及 *RAS* 的扩增在人类肿瘤中占有很大的比例（表 11-2）。

表 11-2 人类恶性肿瘤中扩增的癌基因

肿瘤类型	扩增基因	频率(%)	肿瘤类型	扩增基因	频率(%)
神经母细胞瘤	*MYCN*	20~25	肉瘤	*MDM2*	10~30
小细胞肺癌	*MYC*	15~20		*CDK4*	11
胶质细胞瘤	*EGFR*(*ERBB-1*)	33~50	宫颈癌	*MYC*	25~50
乳腺癌	*MYC*	20	卵巢癌	*MYC*	20~30
	ERBB2	~20		*ERBB2*	15~30
	FGFR1	12		*AKT2*	12
	FGFR2	12	头颈癌	*MYC*	7~10
	CCND1(*cyclin d1*)	15~20		*ERBB2*	10
喉癌	*MYC*	38		*CCND1*(*cyclin d1*)	~50
	CCND1(*cyclin d1*)	25	结肠直肠癌	*MYB*	15~20
胃癌	*KRAS*	10		*HRAS*	29
	CCNE1(*cyclin e1*)	15		*MYCN*	22
肝癌	*CCND1*(*cyclin d1*)	13			

（三）染色体重排

在造血系统恶性肿瘤及实体瘤中经常可检测到染色体重排。这些重排主要是染色体易位，其次是染色体插入。在血液系统中，染色体重排主要通过两种机制致瘤：原癌基因的转录激活或产生融合基因。其中原癌基因转录激活是由于染色体重排导致原癌基因移位至免疫球蛋白或 T 细胞受体基因附近，使原癌基因的转录受免疫球蛋白或 T 细胞受体调节因子的控制，调节原癌基因异常表达并使细胞恶性转化。当染色体断裂点位于两个不同基因的时候，染色体重排就可能形成融合基因，即两个不同的基因片段连接形成一种复合结构，由一个基因的“头”和另一个基因的“尾”组成的融合基因。融合基因编码具有转化活性的融合蛋白。一般来说，参与融合的两个基因均编码具转化能力的融合原癌蛋白。此外，在造血系统恶性肿瘤中，免疫球蛋白或 T 细胞受体基因的重排错误往往又增加了染色体重排发生的频率。

1. 基因激活　在 Burkitt 淋巴瘤中，发现有 75% 的病例存在 t(8;14)(q24;q32)，这是原癌基因转录激活的一个典型例子。染色体重排使位于染色体 8q24 的 *MYC* 基因受到位于 14q32 的免疫球蛋白重链的调节因子控制。易位的结果导致编码调控细胞增殖的核蛋白 *MYC* 基因激活，这在 Burkitt 淋巴瘤中起了关键作用。在 Burkitt 淋巴瘤中，*MYC* 基因也可被与免疫球蛋白轻链基因相关的染色体易位激活。这些染色体易位包括涉及 2p12 上 κ 链的 t(2;8)(p12;q24) 及 22q11 上 λ 链的 t(8;22)(q24;q11)。在 Burkitt 淋巴瘤病例中，尽管与 *MYC* 基因相关的染色体断裂点位置有显著的差异，但易位的结果是相同的，均引起 *MYC* 表达的调节失控，导致细胞异常增殖。

在一些急性 T 细胞性淋巴细胞白血病（T-ALL）中，*MYC* 基因是由 t(8;14)(q24;q11) 易位激活。在这些病例中，*MYC* 基因转录由位于 14q11 的 T 细胞受体基因 α 链上的调节因子调控。

2. 基因融合　基因融合并使原癌基因激活首先是在研究慢性粒细胞白血病中的费城染色体断裂点时发现的。在 CML 中，染色体发生 t(9;22)(q34;q11) 使位于 9q34 的 *ABL1* 基因与位于 22q11 的 *BCR* 基因融合在一起，形成位于 22 号易位染色体上的 *BCR/ABL* 融合基因，它编码分子量为 210kD 的融合蛋白，这种蛋白具有很高的酪氨酸激酶活性。BCR/ABL 融合蛋白可使骨髓肿瘤细胞克隆增

多。在急性淋巴细胞白血病(ALL)中,有20%以上的病例存在t(9;22)易位,其中*BCR*基因的断裂位点不同于CML中的断裂点,这种融合基因编码的BCR/ABL融合蛋白分子量为185kD。基因融合有时也会形成融合转录因子。在儿童ALL中,出现t(1;19)(q23;p13),其结果使*E2A*转录因子基因(19p13)与PBXⅠ蛋白同源区基因(1q23)发生融合。E2A/PBXⅠ融合蛋白由E2A蛋白氨基末端的反向激活区域与PBXⅠ的DNA结合同源区域构成。

第三节 肿瘤抑制基因

肿瘤抑制基因的概念最初是20世纪60年代在肿瘤细胞与正常细胞杂交研究的基础上提出的,当时发现正常细胞与肿瘤细胞融合形成的杂种细胞不具备肿瘤细胞表型,而且正常细胞的染色体可以逆转肿瘤细胞表型。因此,人们提出了正常细胞中存在抑制肿瘤发生的基因,称为肿瘤抑制基因(tumor suppressor genes,TSGs)。肿瘤抑制基因是一类存在于正常细胞中的、与原癌基因共同调控细胞生长和分化的基因,也称抗癌基因(anti-oncogene)和隐性癌基因(recessive oncogene)。肿瘤抑制基因是保护性基因,正常情况下可抑制(限制)细胞生长,此作用通过监控细胞分裂速率、修复错配的DNA(突变诱因)和控制细胞死亡来实现。当肿瘤抑制基因发生突变(由于遗传或环境因素),细胞可持续增长并最终形成肿瘤。自从1986年在人类恶性肿瘤中首次发现肿瘤抑制基因*RB*以来,目前已经发现了20余种肿瘤抑制基因,许多人类遗传型肿瘤综合征常常伴有肿瘤抑制基因的缺失或失活。

一、肿瘤抑制基因的发现

事实上,早在20世纪初期研究癌基因过程中就发现了肿瘤抑制基因存在的线索。德国动物学家Boveri通过海胆卵操作实验(诱导多级有丝分裂和染色体异常分离),发现了有丝分裂纺锤体,并偶然观察到异常的有丝分裂会导致子代染色体缺失,产生与恶性肿瘤中低分化组织团块相像的异常细胞团块,由此推测染色体的不正确联合会产生能够遗传的、有无限增殖能力的恶性细胞。同时,Boveri还用自己的观点解释了和肿瘤相关的许多现象,甚至想象有毒物质、物理损伤、病原体、慢性炎症和组织修复等因素都可能间接促进染色体异常分离或导致染色体不平衡等情况的出现,进而导致肿瘤的发生。除了这些重要的发现与假设,Boveri还阐明了在一种组织中会出现不同类型的肿瘤、隐性染色体等位基因丢失、肿瘤易感性的遗传度等多个问题。然而,Boveri的假说并没有引起重视,首先是由于缺少来自动物及人类核型研究的直接实验证明,而且尚无直接证据证实染色体数目的改变确实导致了肿瘤表型。随后,Charlest和Clausen根据苯并芘致乳头状瘤的实验结果,假设肿瘤的发生可能与细胞中肿瘤抑制基因的失活有关,并提出了肿瘤抑制基因的概念,但也未得到足够重视。而Rous在1910年用患肉瘤鸡的无细胞滤液导致正常鸡患肉瘤的发现,直接证明了肿瘤的遗传基础,引发了癌基因研究的高潮。相比之下,所有肿瘤抑制基因的证据都是非直接的,肿瘤抑制基因的研究迟迟没有进展。

直到1969年,Ephrussi和Harris的体细胞杂交实验才使寻找肿瘤抑制基因的研究拨云见日。他们将小鼠恶性肿瘤细胞与正常小鼠细胞融合后,发现形成的四倍体杂种细胞并无恶性表型,接种到适当宿主体内也不再生长肿瘤。由于四倍体细胞不稳定,传代过程中,来自小鼠正常细胞的染色体会逐渐丢失。理论上来说,随着正常染色体的逐步丢失,杂种细胞的恶性表型逐步恢复(即致瘤回复体)。既然杂种细胞有来自恶性肿瘤细胞中已激活的癌基因,理应致癌,这显然不符合癌基因显性作用方式。由此,Harris等推测小鼠正常细胞中可能存在另一种抑制肿瘤的基因,可以抑制癌细胞的恶性表型,并提出恶性肿瘤是一种隐性性状,在体细胞杂种细胞中因存在来自正常细胞的染色体(基因)而被抑制。随后一系列的啮齿类体细胞杂交实验及啮齿类肿瘤细胞—正常人细胞的杂种细胞也都支持此论点。其后发展的微细胞技术(将单个染色体从正常细胞转移到癌细胞)也证实特定的人正常染

色体可抑制各种癌细胞的致瘤性生长。

与体细胞杂交研究基本同步，Knudson 在进行视网膜母细胞瘤的流行病学研究中发现大多数视网膜母细胞瘤的病例是散发的，但也有以常染色体显性遗传模式传递的家系。与散发型病例相比，遗传型病例发病年龄较早，往往呈双侧或多灶性。1971 年，Knudson 提出了著名的"二次突变或二次打击假说"（two-hit theory）。他认为遗传型与散发型都起源于同一基因的两次以上的突变，一个等位基因的突变不足以形成肿瘤。遗传型视网膜母细胞瘤患者的第一次突变发生于生殖细胞，使婴儿的所有体细胞都含有此突变，成为突变的杂合子，第二次突变则发生于视网膜体细胞。二次突变相加即可完成肿瘤的始动，使良性细胞转变为恶性细胞，因此遗传型病例发病早，常呈双侧。在散发病例中，二次突变均发生在体细胞，而且必须在同一视网膜母细胞中先后发生才能完成肿瘤的始动过程。这种机会很少，需要漫长的时间积累，因此散发病例发病较晚，多为单侧。100 多例视网膜母细胞瘤患者的实际调查结果均符合此假说。此外，Knudson 对其他几种儿童肿瘤（如 Wilms 肾母细胞瘤、多发性内分泌瘤等）的调查分析都支持二次突变理论。

其后的系列研究确定了视网膜母细胞瘤发生的具体染色体和基因座。1976 年 Francke 报道在遗传型视网膜母细胞瘤患者外周血淋巴细胞或皮肤成纤维细胞中都发现 13q14 缺失。1983 年 Cavenee 等用检测杂合性缺失（LOH）方法证实散发型病例瘤体中 13q14 存在杂合性丢失。1985 年 Cavenee 等又在两个不同的视网膜母细胞瘤家族中证明肿瘤中丢失的正是 13 号染色体上的正常等位基因。1987 年李文华等发表了人视网膜母细胞瘤易感基因的克隆、鉴定和序列，这是人类鉴定和分离的第一个肿瘤抑制基因。

二、部分重要的肿瘤抑制基因

（一）*RB* 基因

视网膜母细胞瘤（retinoblastoma，RB）基因是在研究视网膜母细胞瘤家系时发现的，是最早发现的肿瘤抑制基因。

RB 基因位于 13q14.1，转录产生长为 4.7kb 的 mRNA。*RB* 基因共有 27 个外显子。*RB* 基因编码的蛋白质（RB）由 928 个氨基酸组成，分子量为 110kD。RB 是一种核磷酸蛋白质，在视网膜母细胞瘤中均存在缺失，但是在大多数肿瘤细胞中也可以检测到一定量的 RB。

各种不同类型的突变均导致 *RB* 基因功能的丧失。30% 的视网膜母细胞瘤中出现 *RB* 基因的大片段缺失。此外，在一些视网膜母细胞瘤中也发现 *RB* 基因的剪接错误、点突变及启动子区域的微小缺失。

RB 的主要作用是调节细胞周期。其调节细胞周期的能力与 RB 磷酸化状态有关。磷酸化状态在细胞 S 期起始最大，在有丝分裂之后进入 G_1 期时最小。有丝分裂原的刺激可以促进静止细胞中 RB 的磷酸化。相反，细胞分化能够促进 RB 的低磷酸化。因此，RB 的低磷酸化状态可以抑制细胞增殖。在 *RB* 基因参与的生长抑制途径中，一个重要的因子是 *MYC* 基因。TGFB1 诱导的角质形成细胞的增生常伴有 *MYC* 表达的抑制。*MYC* 表达的抑制可以通过导入含 SV40 腺病毒大 T 抗原的表达载体与 RB 结合来拮抗。因此，角质形成细胞中存在 *TGFB1*、*RB*、*MYC* 表达之间的联系。

DNA 肿瘤病毒、SV40、腺病毒、多瘤病毒及人乳头瘤病毒等通过转染进入细胞并与低磷酸化的 RB 结合，从而促进细胞的增殖。

（二）*P53* 基因

P53 基因定位于 17p13.1，长 20kb，含有 11 个外显子，编码 393 个氨基酸组成的分子量为 53kD 的蛋白质。*P53* 基因的突变常发生在结肠癌、乳腺癌、肝癌、肺癌等多种肿瘤中。事实上，与目前已知的任何一种肿瘤抑制基因或癌基因相比，*P53* 基因在 50% 左右的人类恶性肿瘤中存在变异，占第一位。最初人们认为 *P53* 基因是癌基因，但是后来发现某些肿瘤中的 P53 蛋白与正常 *P53* 基因编码的蛋白不同。经过一系列的研究最终证实了 *P53* 基因是一种肿瘤抑制基因。

P53 蛋白含有 4 个主要功能区:①N 端酸性转录激活区,可激活转录,介导蛋白间相互作用。这一区域还可以与 MDM2 蛋白-P53 蛋白的负调控因子结合。②序列特异性结合区(核心区),这一区域具有特异性结合 DNA 的功能,而且是肿瘤细胞中最常检出突变的区域,含有 6 个突变热点,占已知 *P53* 错义突变的 40%。③寡聚区,介导 P53 蛋白自身聚合形成四聚体。④羧基端,可与 DNA 非特异性结合,参与核心区与 DNA 结合的错构调节。

野生型 P53 蛋白是核内的一种磷酸化蛋白质,作为转录因子可与特异的 DNA 序列结合。一些外界刺激,如 DNA 损伤、应激等,可引起细胞内 P53 蛋白水平升高,激活一系列下游靶基因的转录,诱导细胞 G_1 期,阻断细胞凋亡和细胞分化,保护基因组的完整性以及抑制肿瘤细胞的生长等。目前已知的下游靶基因包括 *P21*、*GADD45*、*MDM2*、*BAX*、*IGF-BP3* 和 *PIG3* 等。

P53 基因功能失活机制有以下几种:①*P53* 基因自身突变,导致 P53 蛋白丧失与 DNA 结合的能力,这是 *P53* 基因失活的最重要机制。②*MDM2* 癌基因的负调节:*MDM2* 是 P53 蛋白的靶基因,P53 蛋白刺激 *MDM2* 基因的表达,而 MDM2 蛋白可与 P53 蛋白(野生型或突变型)结合,抑制 P53 蛋白介导的反式激活、增殖抑制和诱导凋亡的功能,同时 MDM2 蛋白可以催化 P53 蛋白的降解,从而形成一个反馈调节环,负调节 P53 蛋白的活性。③P53 蛋白与癌蛋白之间的相互作用可能是其失活的另一重要原因。DNA 肿瘤病毒蛋白,如 SV40 大 T 抗原、腺病毒 E1B 转化蛋白、人乳头瘤病毒 E6 蛋白等,均可以和 P53 蛋白结合,抑制其功能活性并促进其降解。

由于 *P53* 基因在肿瘤发生、发展以及诊断治疗中的重要作用,目前科学家正致力于寻找和发现其相关基因以及应用于基因治疗的有效方法。

(三) *WT 1* 基因

Wilms 瘤是儿童常见的一种肾脏恶性肿瘤,累及单侧或双侧肾脏。遗传学研究表明,染色体上至少有 3 个位点与 Wilms 瘤发生相关。散发性 Wilms 瘤常伴有染色体缺失,缺失区集中在 11p13 和 11p15。1990 年,通过对 Wilms 瘤及其导致的 WAGR 综合征患者进行分析,确认其中的一个位点是 *WT1*,称为 Wilms 瘤候选基因 1,位于 11p13,全长 345kb,含有 10 个外显子,转录的 mRNA 长 3kb,其转录本存在选择性剪接。*WT1* 基因编码的蛋白在进化过程中高度保守,提示此基因具有重要作用。WT1 主要在胚胎肾组织中高表达。WT1 蛋白含有几个功能结构域,其中包括 4 个锌指结构域,提示它是个转录因子;蛋白中心有亮氨酸拉链,提示 WT1 可能与其他具有亮氨酸拉链的蛋白质相互作用;N 端的 180 个氨基酸与自我结合有关。WT1 蛋白可以抑制一些生长诱导基因启动子元件的转录活性,这些基因包括早期生长受体(*EGR1*)、胰岛素样生长因子 2(*IGF-2*)和血小板源生长因子的 A 链(*PDGFA*)基因,这提示 *WT1* 是一种肿瘤抑制基因。

与 *RB* 基因不同,*WT1* 基因的表达具有组织局限性。11p13 区域纯合性缺失的 Wilms 瘤中检测不到 *WT1* 基因的 mRNA。另有一些 Wilms 瘤表现为 *WT1* mRNA 高表达,这是由突变的 *WT1* 基因产生的。

(四) *CDKN 2A* 基因

细胞周期蛋白依赖性激酶抑制剂 2A 基因(Cyclin Dependent Kinase Inhibitor 2A,*CDKN2A*),又称多重肿瘤抑制基因(multiple tumor suppressor 1,*MTS1*),由于其编码的蛋白产物分子量为 16kD,也被称为 *P16* 基因。Kamb 等和 Nobori 等在研究与细胞周期依赖性激酶 4(cell cycle dependent kinase 4,*CDK4*)作用的蛋白时,几乎同时发现了 *P16* 基因。*P16* 基因是细胞周期依赖性激酶 4 的抑制因子(cell cycle dependent kinase 4 inhibitor,*CDK4I*),定位于 9p21,全长 8.5kb,含有 3 个外显子,编码 148 个氨基酸组成的 16kD 蛋白质。*P16* 基因的缺失与突变见于各种恶性肿瘤,包括黑色素瘤、胶质瘤、肺癌、白血病等多种肿瘤。

P16 基因蛋白产物的缺失主要与 *P16* 基因本身的缺失、点突变及重组有关,其中 *P16* 基因的纯合性缺失是主要原因,*P16* 基因启动子区域的突变也可以导致 *P16* 基因失活,转录与翻译水平上的调控机制异常也可使 *P16* 基因功能产物丧失。此外,*P16* 基因 5′端 CpG 岛的甲基化可导致

P16 基因转录停滞。

P16 基因主要是通过调节细胞周期来参与肿瘤发生，此调节机制是通过 *RB* 基因依赖性途径实现的。P16 蛋白能抑制 CDK4/CDK6 介导的 RB 磷酸化。未磷酸化的 RB 蛋白可以结合到转录因子 E2F-DNA 复合物上，在启动子水平阻断 mRNA 的转录，同时又可活化转录因子 ATF2，促进细胞分裂抑制因子的转录，磷酸化的 RB 蛋白无此活性。G_1 期 cyclinD-CDK4 复合物具有激酶活性，促使 RB 磷酸化，而 P16 蛋白与 cyclinD 竞争性结合 CDK4，从而抑制 RB 磷酸化。如果 *P16* 基因发生突变或缺失，就解除了对 cyclinD-CDK4 复合物的抑制，使其发挥抑制 RB 的功能，细胞周期中的 G_1 关卡出现异常，细胞获得无限增殖能力。

（五）*CDKN 2B* 基因

在研究 *P16* 基因的同时，发现 9p21 区域部分序列发生突变，而 *P16* 基因又能保持完整，由此推断此区域可能存在多种肿瘤抑制基因。经过基因克隆和序列分析发现了另一个多重肿瘤抑制基因，称为细胞周期蛋白依赖性激酶抑制剂 2B 基因（Cyclin Dependent Kinase Inhibitor 2B，*CDKN2B*），也称 *MTS2*，由于其编码分子量为 15kD 的蛋白质，又称 *P15* 基因。*P15* 基因的 cDNA 全长 837bp，编码由 137 个氨基酸组成的蛋白质。*P15* 基因与 *P16* 基因具有高度的同源性，从功能上来说，*P15* 基因与 *P16* 基因也具有相似性。P15 蛋白可以特异性抑制 cyclinD-CDK4、cyclinD-CDK6 复合物的蛋白激酶活性。

P15 基因作为 *P16* 基因家族中的第二个成员，对多种恶性肿瘤均具有抑制作用。*P15* 基因突变与一些肿瘤发生之间有一定的关联，如儿童急性淋巴母细胞性白血病和非小细胞性肺癌。

（六）*NF1* 基因

1989 年，Fountain 等和 O'Connel 等通过连锁分析将 *NF1* 基因定位在 17 号染色体，它是在研究神经纤维瘤（neurofibromatosis）时被发现的。*NF1* 基因的 mRNA 长约 11～13kb，包含 7.5kb 的编码区。编码的蛋白质（NF1）含有 2 485 个氨基酸，与 *RAS* 高度同源。NF1 蛋白又称为神经纤维瘤蛋白（neurofibromin），它在全部组织中均有表达。

良性神经纤维瘤与恶性神经纤维瘤发生的不同之处在于是 1 个 *NF1* 等位基因还是 2 个 *NF1* 等位基因的失活。然而，在神经纤维肉瘤中，除了 *NF1* 基因失活外，还发现有其他的遗传学改变。在神经纤维肉瘤中均可见 17p 缺失，而在神经纤维瘤中无此改变。

（七）*NF2* 基因

2 型神经纤维瘤（*NF2*）是一种常染色体显性疾病，也称为中心型神经纤维瘤，从遗传背景到临床特征都与 *NF1* 明显不同。*NF2* 的特点是发生于双侧的雪旺氏细胞瘤，使第八对颅神经前庭支受累（听神经瘤）。*NF2* 患者伴发脑膜瘤、脊柱雪旺氏细胞瘤和室管膜瘤的风险也增高。通过连锁分析和 LOH 分析将 *NF2* 基因定位于 22q，并于 1993 年通过定位克隆策略成功克隆到 *NF2* 基因。在 *NF2* 患者中可检测到 *NF2* 基因的生殖细胞突变，*NF2* 基因的体细胞突变也见于散发性雪旺氏细胞瘤和脑膜瘤。在其他类型的肿瘤中，*NF2* 基因的体细胞突变并不常见。*NF2* 基因编码蛋白结构与细胞骨架蛋白家族高度同源，是一种内膜蛋白和结构蛋白之间的连接成分。由此可见，*NF2* 基因的变异通过影响细胞形态、细胞间相互作用和细胞的运动参与肿瘤的演进。

（八）*CDKN 1A* 基因

细胞周期蛋白依赖性激酶抑制剂 1A 基因（Cyclin Dependent Kinase Inhibitor 1A，*CDKN1A*）是 1993 年发现并克隆的。*CDKN1A* 基因定位于 6p21.2，全长 2.1kb，编码 164 个氨基酸组成的蛋白质，分子量为 21kD，所以又称 *P21* 基因。P21 蛋白是 P53 作用的靶蛋白。P21 可以通过抑制多种细胞周期蛋白与 CDK 的复合物和增殖细胞核抗原的活性而直接作用在 S 期前或 S 期，阻止 DNA 的复制，抑制细胞增殖。

（九）*CDKN 1B* 基因

细胞周期蛋白依赖性激酶抑制剂 1B 基因（Cyclin Dependent Kinase Inhibitor 1B，*CDKN1B*）是

1994 年被克隆的。*CDKN1B* 基因定位于 12p13，cDNA 长度为 594bp，编码 198 个氨基酸组成的蛋白质，分子量为 27kD，所以也称 *P27* 基因。*P27* 与 *P21* 属于同一基因家族，在氨基酸水平上，两者具有明显的同源性。

P27 主要作为 CDK 抑制因子参与细胞周期调控。正常情况下，P27 在 G_0/G_1 期表达增高，当细胞进入 S 期时表达下降。P27 可以抑制细胞的 DNA 合成，阻止细胞从 G_1 期进入 S 期。P27 是一个非特异性 CDK 抑制因子，P27 能与 cyclinD-CDK4、cyclinD-CDK6、cyclinE-CDK2、cyclinA-CDK2 等复合物结合，抑制其活性，使细胞停滞于 G_1 期。此外，P27 还具有启动细胞凋亡的作用。

P27 的表达与许多肿瘤的发生及发展相关，P27 低表达可作为恶性肿瘤预后不良的一个独立指标。目前，许多学者正在尝试应用 *P27* 基因进行肿瘤细胞的基因治疗。

（十）*BRCA1* 和 *BRCA2* 基因

家族史一直是乳腺癌发病的一个重要风险因素，一级亲属中有多人患乳腺癌的个体发病风险最高。20 世纪 80 年代末，研究证实家族性乳腺癌易感性与一种高外显率的常染色体显性等位基因相关。1990 年，Hall 及其合作者报道了一个乳腺癌易感基因，称为 *BRCA1*（breast cancer predisposition gene 1），*BRCA1* 定位于 17q21，表达序列长 7.8kb，含有 22 个外显子，编码 1 863 个氨基酸组成的蛋白质，其 N 端附近含有锌指结构域，包括半胱氨酸和组氨酸残基构成的保守序列。*BRCA1* 基因是一个乳腺与卵巢组织特异性肿瘤抑制基因，它符合 Knudson 的二次突变学说。*BRCA1* 基因的种系突变不但能大大增加乳腺癌的发病风险，而且可以增加卵巢癌的发病风险。在家族性乳腺癌中，*BRCA1* 基因在生殖细胞中已经发生了一次突变，当乳腺组织再次发生突变时，形成杂合性缺失（loss of heterozygosity，LOH），导致肿瘤的发生。

BRCA1 基因的种系突变可作为 40% ~50% 乳腺癌家系肿瘤易感性的主要原因，而在那些无 *BRCA1* 基因突变的乳腺癌家系中，发现了另外一个具有高外显率的常染色体显性易感基因 *BRCA2*。*BRCA2* 基因定位于 13q12-q13，mRNA 长 11kb，编码 3 418 个氨基酸组成的蛋白质。*BRCA2* 和 *BRCA1* 基因种系突变涉及约 80% 乳腺癌的发病风险，但对于卵巢癌而言，由 *BRCA2* 基因突变所导致的发病风险为 10%，而 *BRCA1* 基因则为 40% ~50%。携带 *BRCA2* 基因突变的男性患乳腺癌的风险大大提高，约为 6%。相反，携带 *BRCA1* 基因突变的男性却没有患乳腺癌的风险。无论男性还是女性，携带 *BRCA2* 基因突变的个体发生胰腺癌及其他肿瘤的风险都有所提高。

（十一）*DCC* 基因

18 号染色体的杂合性缺失常见于结直肠肿瘤（73%）和恶性腺瘤（47%），但仅见于少数早期的腺瘤（11% ~13%）中。1990 年 Fearon 等发现 18 号染色体的缺失区域在 18q21.2 到末端。在此区域克隆到了 370kb、包含有肿瘤抑制基因的 DNA，用表达的外显子作探针筛查 cDNA 文库，最后获得了编码基因，该基因由于在结肠直肠癌中存在缺失而称为 *DCC*（deleted in colorectal carcinomas）基因。*DCC* 全长约 1.4Mb，含有 29 个外显子。

DCC 基因在结肠腺体等正常组织中表达。结直肠癌中可见 *DCC* 基因的体细胞突变。突变类型包括 5′端的纯合性缺失、内含子的点突变和外显子下游 170bp 片段的 DNA 插入等。

DCC 蛋白是一种跨膜蛋白，属于免疫球蛋白家族，分子结构与细胞黏附因子包括神经细胞黏附因子（N-CAM）相似。

DCC 基因可能通过在肿瘤生长和转移过程中诱导细胞凋亡来发挥其肿瘤抑制基因的作用。

（十二）*APC* 基因

遗传型结肠直肠癌综合征通常分为息肉型和非息肉型两种。息肉型在恶性肿瘤发生之前可见成百上千的良性肿瘤（息肉），息肉有着很高的恶变风险，多数结肠直肠癌都起源于腺瘤性癌前病变，但这种息肉在非息肉型结肠直肠癌却十分罕见。遗传型结肠直肠癌综合征中有一种息肉病综合征称为家族性结肠息肉病（FAP）或腺瘤性结肠息肉病（APC）。APC 是一种常染色体显性遗传病，在家族性

结肠息肉癌变过程中,一个*APC*等位基因的丢失发生在结肠腺瘤阶段,另一个正常的等位基因丢失发生在结肠腺瘤向腺癌转变的阶段。该病以患者早期出现发生于结肠和直肠的上百个腺瘤性息肉为特征。具有典型家族性结肠息肉的个体其结肠直肠癌的发病风险极高,到40岁时接近100%出现恶变。

1986年,Herrera和Sandberg对一个具有典型家族性结肠息肉症状但无家族史患者的染色体进行分析,发现其5号染色体长臂存在缺失,为*APC*基因定位提供了最早的依据。随后对大量家族性结肠息肉病家系进行DNA连锁分析将*APC*基因定位于5q21。*APC*基因含有15个外显子,全长125kb,mRNA长8.5kb,编码2 844个氨基酸的蛋白质。APC蛋白分子量为300kD,定位于结肠上皮细胞基膜侧,其功能涉及调节细胞黏附、细胞迁移乃至细胞凋亡。

大约80%家族性结肠息肉病或Gardner综合征患者有*APC*基因的杂合性种系突变,这些突变会使APC蛋白失活。绝大多数突变为位于*APC*基因5′端编码区的无义突变或移码突变。携带*APC*种系突变的个体一旦发生另一个野生型等位基因的体细胞突变就会产生肿瘤。

(十三) *P73*基因

*P73*基因是1997年由Kaghad等在研究胰岛素信号介质时,对COS细胞的cDNA文库进行杂交筛选时发现的。近年来发现*P73*基因与*P53*基因在结构与功能上非常相似,属于*P53*基因家族。它定位于1p36,含有13个外显子,编码分子量为73kD的蛋白质。P73具有与P53同样的生长抑制和促凋亡效应,还具有与*P53*不同的单基因表达方式,不与病毒蛋白相互作用。*P73*与肿瘤起源的关系密切,但是其在肿瘤发生中的具体作用机制尚不清楚。

(十四) *VHL*基因

Von-Hippel-Lindau(VHL)综合征是一种罕见的显性遗传疾病,主要引起中枢神经系统和视网膜的血管母细胞瘤,以及透明细胞型(clear cell type)肾癌和嗜铬细胞瘤。通过连锁分析将*VHL*基因定位于3p。LOH研究显示该基因具有肿瘤抑制基因特性,即在肿瘤发生过程中两个等位基因全部失活。1993年通过定位克隆确认了*VHL*基因。在多数VHL综合征患者中可检测到*VHL*基因的生殖细胞突变,使一个*VHL*等位基因失活。值得注意的是,某一种类型的*VHL*生殖细胞突变只参与肾癌的发生,另一种突变与肾癌和嗜铬细胞瘤的发生均相关,还有的突变只与嗜铬细胞瘤的发生相关。*VHL*基因体细胞突变在散发性透明细胞型肾癌中的检出率可达80%,但在其他组织类型的肾细胞癌中却未检出。

*VHL*基因编码由213个氨基酸组成的蛋白质,它与SⅢ转录延伸因子B、C亚单位形成复合体,调节RNA聚合酶Ⅱ的转录延伸过程。此外,VHL和延伸因子B、C在胞浆中可形成多种蛋白复合体,其中一种复合体的作用是降解细胞中的其他蛋白质。

(十五) *NM 23*基因

*NM23*基因是1988年由Steeg首先从黑色素瘤K-1735细胞系中通过消减杂交(subtractive hybridization)方法克隆到的能抑制肿瘤细胞转移的基因。*NM23*基因编码由153个氨基酸组成的分子量为17kD的蛋白质。*NM23*基因家族有两个成员:*NM23-H1*和*NM23-H2*,均定位于17q21.3,二者有高度的同源性。NM23蛋白具有核苷二磷酸激酶的活性,还有嘌呤结合功能。*NM23*基因是一种肿瘤转移抑制基因(tumor metastasis suppressor gene),表达水平在低转移性肿瘤中明显高于高转移性肿瘤。将*NM23*基因转染(transfection)到高转移肿瘤细胞中,可使癌细胞转移潜能下降。目前发现*NM23*基因参与乳腺癌、肺癌、结肠癌等多种恶性肿瘤的转移过程。

第四节 肿瘤发生的遗传学理论

一、肿瘤发生的单克隆起源假说

致癌因子引起体细胞基因突变,使正常体细胞转化为前癌细胞,然后在一些促癌因素作用下,发

展成为肿瘤细胞。按照这个学说的观点，肿瘤细胞是由单个突变细胞增殖而成的，也就是说肿瘤是突变细胞的单克隆增殖细胞群，这称为肿瘤的单克隆起源学说。肿瘤的细胞遗传学研究结果证实，所有的肿瘤几乎都是单克隆起源，也就是说患者的所有肿瘤细胞都起源于一个前体细胞。最初是一个关键的基因突变或一系列相关事件导致单一细胞向肿瘤细胞的转化（transformation），随后产生不可控制的细胞增殖，最后形成肿瘤。

有许多证据可证明肿瘤的克隆特性，其中包括白血病和淋巴瘤的分子分析，表明所有的淋巴瘤细胞都有相同的免疫球蛋白基因或T细胞受体基因重排，提示它们来源于单一起源的B细胞或T细胞。而体细胞突变（somatic mutation）和克隆选择（clonal selection）模式说明肿瘤在构成上是单克隆的，女性X连锁基因的分析提供了肿瘤克隆特性的最初证据。尽管女性的所有细胞都包含两条X染色体，但在早期胚胎形成时其中有一条随机失活。因此每一位女性在细胞构成上来说是嵌合的，一部分细胞中为其中一条X染色体失活，另一些细胞中则是另外一条X染色体失活。如果一条X染色体上的基因与另一条X染色体上的等位基因不同，就可以区分这两种细胞。葡萄糖-6-磷酸脱氢酶（G6PD）基因是一个X连锁基因，在部分人群中存在高突变率，杂合子（heterozygote）个体一条X染色体上有一个野生型基因（wild-type gene），另外一条X染色体上相应等位基因失活。失活的X染色体可以通过依赖于G6PD活性的细胞染色得到验证。这种情况下，正常组织是包含有活性的和失活的G6PD细胞的嵌合体。对女性肿瘤的研究发现，一些恶性肿瘤的所有癌细胞都含有相同失活的X染色体，表明它们是单一细胞起源。

同时，肿瘤细胞学研究发现同一肿瘤中所有肿瘤细胞都具有相同的标记染色体，再次证明恶性细胞的单克隆起源。近年来，通过FISH方法对癌组织中突变的癌基因或肿瘤抑制基因进行直接分析也证实了肿瘤的单克隆特性。

二、肿瘤发生的癌基因理论

癌基因理论认为，人类肿瘤的发生、发展与体细胞中累积的各种遗传物质改变相关。这些遗传变异涉及染色体重排和癌基因的激活，这对于肿瘤的发生至关重要。癌基因是原癌基因的突变形式，而原癌基因则是一组与细胞生长调控有关的基因，包括生长因子、生长因子受体、信号转导因子、转录因子及程序性细胞死亡调节因子。原癌基因可因基因突变、染色体重排或基因扩增而被激活。原癌基因原始功能的较小改变即可使其成为癌基因。一方面，原癌基因的突变导致蛋白结构的改变，从而出现蛋白（酶）活性提高和调节功能的丧失；另一方面，原癌基因的突变导致蛋白丰度的提高，主要由于蛋白表达提高（通过错误调控）、蛋白稳定性提高和基因的重复（染色体异常的一种），导致细胞中蛋白量的提高。此外，染色体易位导致在错误的时间或者错误的细胞中基因出现高表达，并产生有活性的融合蛋白。近来发现microRNA的突变也可以导致癌基因的激活。正常情况下mircoRNA可以通过下调癌基因的表达而起到调控作用，所以反义RNA可以起到抑制癌基因的作用。可见，基因的改变是肿瘤起源与发展的分子基础。一旦这些基因在表达时间、表达部位、表达数量及表达产物结构等方面发生了异常，就可以导致细胞无限增殖并出现恶性转化。

癌基因的发现为研究肿瘤的分子及遗传学基础提供了突破点。癌基因也为我们理解正常细胞的增殖、分化及程序性死亡提供了重要的线索。异常癌基因的确认为肿瘤的分子诊断及监测提供了工具。更重要的是，癌基因代表了新一代肿瘤治疗的靶标，应用针对特异癌细胞靶点的新型化疗制剂治疗肿瘤已经不仅仅是一个梦想。在未来，新的抗肿瘤药物可以选择性杀死肿瘤细胞，而保留正常细胞。

三、肿瘤发生的肿瘤抑制基因理论——Knudson的二次突变假说

20世纪70年代Alfred Knudson提出了肿瘤抑制基因模式，以解释遗传型视网膜母细胞瘤的发病机制。他在分析同一种类肿瘤遗传型和非遗传型肿瘤之间关系时，专门研究了发生双侧（遗传型）和

单侧(非遗传型)视网膜母细胞瘤的特征,提出上述两种类型的视网膜母细胞瘤之间存在联系,并假设两种类型视网膜母细胞瘤都是由两个独立与连续的基因突变产生的,即二次突变事件引起的。遗传型肿瘤病例中,第一次突变发生于生殖细胞,是种系突变,而第二次突变随机发生在体细胞中。在这种情况下,双侧视网膜的细胞都有可能发生第二次突变并形成肿瘤。相比之下,非遗传型视网膜母细胞瘤是同一个体细胞发生两次独立的突变,因而在双侧视网膜都发生二次突变的可能性较小。Knudson 提出了二次突变假说,它能最简单地解释肿瘤的显性特征(图 11-6)。二次突变假说还表明肿瘤的发生是一种隐性事件,即野生型基因产物可以抑制肿瘤产生,而肿瘤中的这一对等位基因发生了失活。Knudson 称这种基因为抗癌基因,也称肿瘤抑制基因。二次突变假说得到多数遗传型视网膜母细胞瘤患者的 G 显带分析结果的支持,即肿瘤细胞有 13q14 缺失。在克隆出 *RB* 基因后,应用突变扫描分析发现视网膜母细胞瘤的 *RB* 基因均发生了突变或缺失。Knudson 完成了理解肿瘤抑制基因概念的框架性工作。

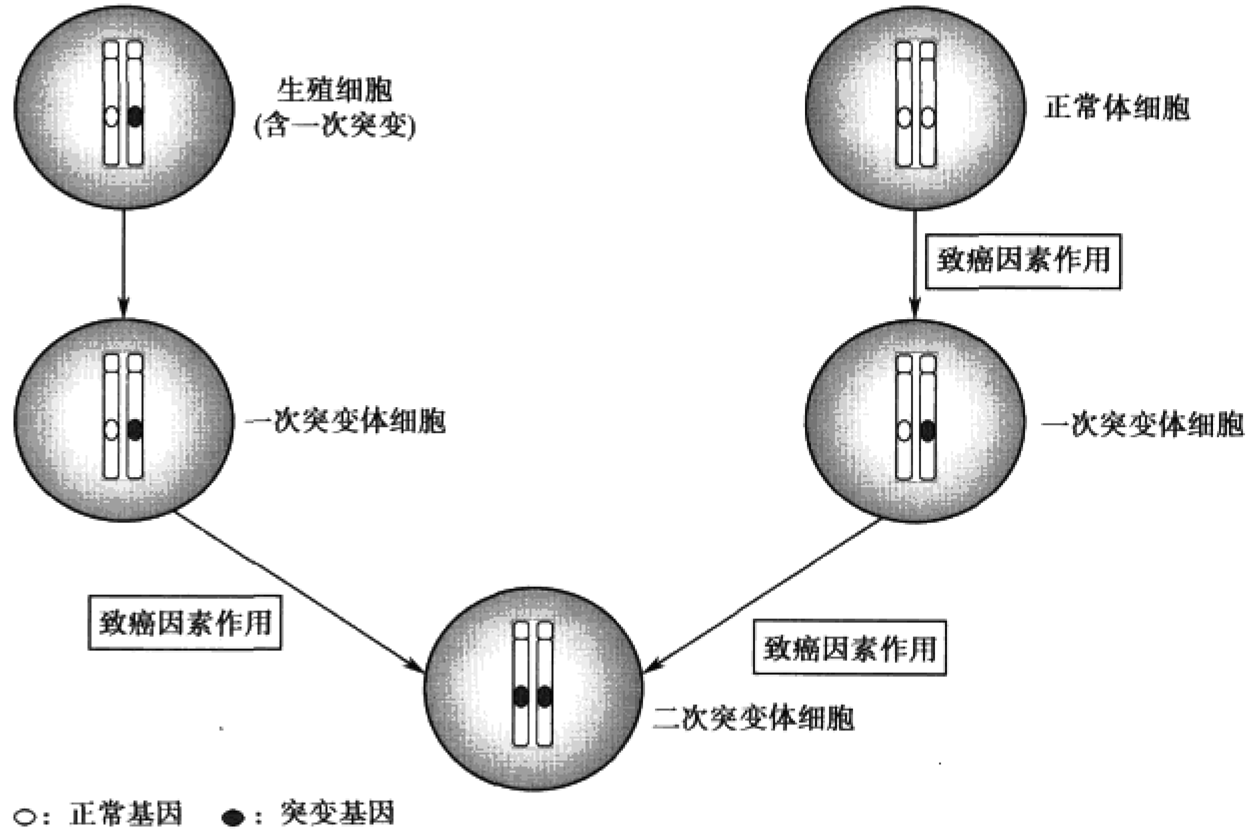


图 11-6 Knudson 二次突变假说示意图

20 世纪 80 年代的一系列实验结果更证明了 Knudson 二次突变假说的精确性,并认为需要两次或两次以上基因突变才能形成恶性肿瘤,而两次突变所发生的时期或阶段就决定了是否是遗传型。遗传型肿瘤是由于第一次突变发生在生殖细胞或者是由于父母遗传而来,所以该个体的所有体细胞实质都是潜在的前癌细胞,任何体细胞如果发生第二次突变就会转化为肿瘤细胞,因此这种肿瘤发生具有家族性、多发性(双侧性)和早发性的特点。而非遗传型肿瘤则是由于第一次突变发生在某个体细胞中,只影响这个体细胞增殖而来的细胞克隆,成为前癌细胞,如果这个体细胞或其克隆发生第二次突变才可形成肿瘤。因此,非遗传型肿瘤发病迟,并具有散发性(单发性)和单侧性等特点。这一学说除了被用于分析视网膜母细胞瘤外,还解释了 Wilms 瘤等儿童肿瘤的发病原因和规律,而后被广泛用于分析各种肿瘤的发生。但该学说对于肿瘤发生中的各种遗传因素和环境因素的影响解释粗糙,不能具体分析,这是该学说的弱点。

但是目前,Knudson 二次突变假说也遭遇到挑战。有研究发现某些肿瘤抑制基因仅有一个拷贝失活即可有效引发肿瘤,此现象称为单倍剂量不足(haploinsufficiency),是指一个等位基因突变后,另一个等位基因能正常表达,但这只有正常水平 50% 的蛋白质不足以维持细胞正常的生理功能。大量

实验证实，由单倍剂量不足引发肿瘤的时间要长于经二次突变的致癌过程。

四、肿瘤发生的染色体理论

20 世纪初，Boveri 发现两次受精的海胆幼胚细胞呈不均等分裂，染色体分配不平衡，这种细胞与肿瘤细胞相似，失去正常生长特点。Boveri 在总结肿瘤细胞中呈现的一些特殊现象的基础上，提出了肿瘤的染色体理论，这个理论认为肿瘤细胞来源于正常细胞，具有某种异常染色体的细胞是一种有缺陷的细胞，染色体畸变是引起正常细胞向恶性转化的主要原因。但是，由于动物细胞染色体的制备和分析技术的限制，Boveri 假说直到 40 年后才被证实。

除血液系统恶性肿瘤外，我们很少在同一种实体瘤中观察到高度一致的"恒定的或特异的"标记染色体，这可能归结于大多数实体瘤的遗传复杂性和不稳定性，如 Shapiro 等在克隆分析胶质细胞瘤亚群时所发现的肿瘤间或肿瘤内的高度异质性。

染色体数目或结构改变可能导致不同的分子事件发生，包括基因的激活、失活、转录调节异常、扩增、缺失，并导致基因及相关区域的结构改变。这些变化可能涉及癌基因或肿瘤抑制基因序列，其他还可能涉及代谢途径控制、组织特异性分化调节以及编码生长因子或细胞-细胞相互作用相关的表面膜分子等基因序列，通过改变细胞的生长与分化并使受累细胞克隆瘤性增殖。从血液系统恶性肿瘤所获研究结果来看，可以假定，在不同肿瘤中所见潜在转化序列（原癌基因）的结构和功能的改变可能是受染色体变化的影响。遗憾的是，除了神经母细胞瘤细胞的 HSRs 和 DMs 可见 *MYCN* 基因扩增，以及与 7 号染色体的重复和结构改变相关的上皮生长因子受体 *EGFR* 基因（基因定位于 7p12. 3-p12. 1）的过表达外，关于染色体介导的原癌基因改变在人类实体瘤中的作用尚缺乏明确的证据。与此相反，许多研究证实了正常染色体对恶性细胞表型的抑制作用。对人种内体细胞杂种的转化和肿瘤表型遗传控制的分析表明：染色体不平衡与肿瘤发生相关；具有相对过剩的、肿瘤来源的染色体，其杂种细胞具备转化和肿瘤特性。种内杂交（正常二倍体/人类恶性细胞）实验证实 1 号及 11 号染色体对肿瘤形成具有抑制作用。Wilms 瘤细胞中通过微细胞融合导入正常的 11 号染色体可以抑制细胞的肿瘤特性。染色体介导的肿瘤抑制作用表明有些基因（肿瘤抑制基因或肿瘤相关基因）显著阻抑细胞水平的恶性表型表达，这些基因的纯合丢失或失活对肿瘤的形成产生了重要作用。

大量或有限的、渐进性获得的累积染色体改变可能对肿瘤克隆的进展起促进作用。早、晚期或转移性的纤维母细胞瘤、黑素瘤和膀胱癌中可见到不同的染色体异常。早期神经母细胞瘤的特征为众数在三倍体范围，结构异常少见，DMs 或 HSR 缺如。与此相反，进展期或复发的神经母细胞瘤中染色体众数在二倍体或亚四倍体范围，并可见到复杂的结构变化及与 *MYCN* 基因扩增有关的频发 DMs 和 HSR。

五、肿瘤的多步骤遗传损伤学说

目前的研究证明肿瘤的发生是多步骤的，涉及多种相关基因包括癌基因和抑癌基因的变异。一种肿瘤会有多种基因的变化，而同一种基因的改变也会在不同种类肿瘤的发生中起作用，大多数肿瘤的发生与癌基因的活化和（或）肿瘤抑制基因的失活有关。美国麻省理工学院的 Land 等人在 1983 年发现，若只用 *EJ-HA-RAS* 癌基因转染，仅能诱导体外培养的大鼠胚胎成纤维细胞发生过量增殖，但并未出现癌变；然而若将 *RAS* 癌基因与 *MYC* 病毒癌基因共同转染，则能使这些细胞转化为癌细胞。由此可见，细胞的癌变至少需要两种致癌基因的联合作用，每一个基因的改变只完成其中的一个步骤，另一些基因的变异最终完成癌变过程。随后，这个观点得到了许多实验结果的进一步证实，并逐渐发展成为人们普遍认同的多步骤致癌（multistep carcinogenesis）假说，也称多步骤损伤学说。目前认为，恶性肿瘤的发生是一个多阶段逐步演变的过程（图 11-7），肿瘤细胞是

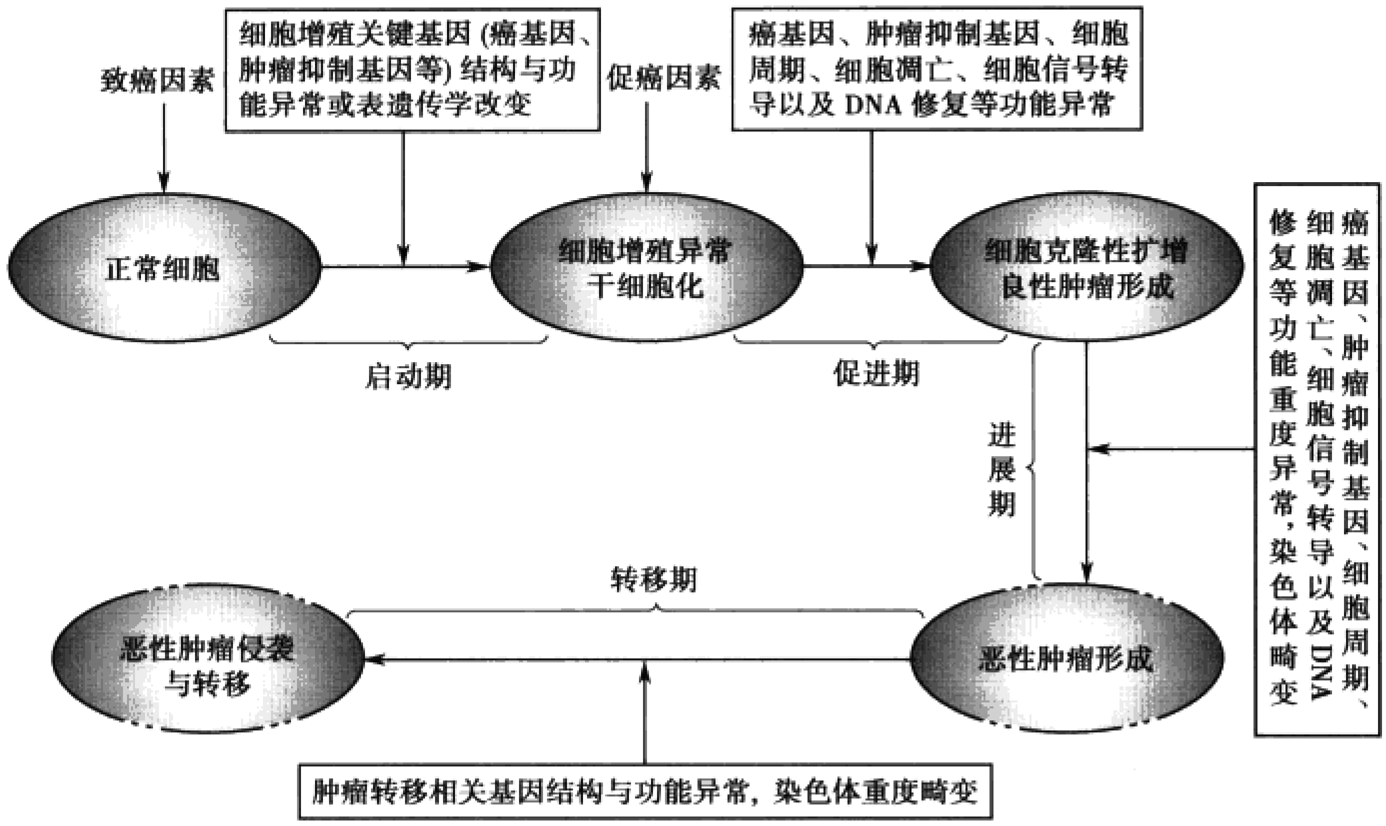


图 11-7 恶性肿瘤多步骤发生机制

通过一系列进行性的改变而逐渐变成恶性的。在这种克隆性演化过程中，常积累一系列的基因突变，可涉及不同染色体上多种基因的变化，包括癌基因、肿瘤抑制基因、细胞周期调节基因、细胞凋亡基因及维持细胞基因组稳定性的基因(包括：DNA 修复、DNA 复制及染色体分离基因)等。这些基因的变化，有的是从种系细胞(germline cells)遗传而来，有的则是从体细胞(somatic cells)由环境因素引起而后天获得的，故肿瘤有遗传型和散发型之别。在肿瘤进展过程中，肿瘤细胞群中常有其他的基因突变发生，授予细胞选择性优势，例如更快速的生长，或具有侵犯和转移的特性，使它们在肿瘤细胞群中占据优势，该过程称为克隆性选择。通过克隆性选择，肿瘤生长更加快速和恶性表型增加。在多步骤损伤学说的基础上，目前将致癌过程分为 3 个阶段：启动(initiation)期，促进(promotion)期和进展(progression)期。

第五节 遗传型恶性肿瘤

部分人类恶性肿瘤的发生具有家族聚集性，也就是一个家族内有多个成员患有同一种肿瘤或几种肿瘤，这种现象又称为遗传型恶性肿瘤综合征。随着人类恶性肿瘤分子遗传学的研究进展，我们对其发生的分子机制有了更深刻的认识。目前认为，各种癌基因、肿瘤抑制基因、生长相关基因、细胞周期调控基因、信号转导基因、细胞凋亡相关基因等的改变均是肿瘤发生的遗传基础，它们构成了个体对肿瘤的遗传易感性。这种对肿瘤的遗传易感性可以从亲代传递到子代，使子代更易患肿瘤，并对环境致癌因素更加敏感。遗传型恶性肿瘤综合征在人群中具有发病早、恶性程度高和多发性等特点，符合孟德尔遗传规律。

一、常染色体显性遗传的恶性肿瘤综合征

(一) 视网膜母细胞瘤(MIM 180200)

视网膜母细胞瘤(retinoblastoma，RB)是一种起源于胚胎视网膜细胞的眼内恶性肿瘤，遗传方式为常染色体显性遗传，儿童早期即发病，发生率约 1/21 000 ~ 1/10 000，多在 4 岁以前发病。其临床表现为早期出现眼底灰白色肿块，多无自觉症状，所以很难发现。此后，肿瘤长入玻璃体，使瞳孔呈黄

色光反射时，才易被发现，称为“猫眼”。

具有染色体 13q 中间缺失的视网膜母细胞瘤的患者均有典型的面貌特征：前额突出、鼻根低且宽、鼻短呈球状、嘴大、上唇薄及耳垂突出。

视网膜母细胞瘤可分遗传型和非遗传型两类。遗传型为双侧发病，约占全部病例的 20% ~ 25%，多在 1 岁半以前发病，可见有家族史。如双亲正常，则双亲之一可能（10%）为携带者。因此，遗传方式符合常染色体显性遗传。

非遗传型多为单侧发病，且在 2 岁以后才发病，约占全部病例的 75% ~ 80%。10% 的单侧发病病例为遗传型的，其双亲之一可能为携带者或患者为新生的突变所致。

遗传型病例中，第一次突变发生于生殖细胞中的 *RB* 基因，结果每一个视网膜母细胞均带有一个突变，成为突变的杂合子。在这个基础上发生的第二次 *RB* 基因突变是体细胞突变（somatic mutation）。二次突变相加，即可完成启动（initiation）而从良性细胞变成恶性细胞。

在非遗传型病例中，二次突变都是体细胞突变，而且必须在同一视网膜母细胞中先后发生才能完成始动过程。这种机会很少，需要经过漫长过程的积累。因此，非遗传型视网膜母细胞瘤多为单发，且发病较晚。

（二） Wilms 瘤（MIM 194070）

Wilms 瘤（Wilms tumor，WT）或称肾母细胞瘤（nephroblastoma），是一种婴幼儿常染色体显性遗传的恶性胚胎肿瘤。本病的名称来自人名，即德国的外科医师 Max Wilms 博士（1867 ~ 1918），是他首先报道了这种肿瘤。Wilms 瘤是由于保留有胚胎分化潜能的肾干细胞功能异常所致，发病率约为 1/10 000，占婴儿肿瘤的 8%，3/4 的肿瘤发生在 4 岁前，90% 在 20 岁内发生。可分为遗传型和非遗传型，遗传型多为双侧发病且发病早，符合常染色体显性遗传，约占 38%；非遗传型常为单侧且发病年龄晚，约占 62%。患者腹部有无症状的肿块，肿块光滑、质坚硬。

Wilms 瘤在很多先天性形态异常综合征中发病率增高，但是这种合并发病的情况仅占临床总发病率的 5%。Wilms 瘤临床表现如伴有无虹膜、偏身肥大、泌尿生殖系统异常和智力低下，称为 WAGR 综合征，该综合征患者有 11 号染色体短臂的中间缺失，即 del(11)(p13)。

研究表明，Wilms 瘤基因是一种肿瘤抑制基因，目前发现与 WT 发生相关的基因及染色体区域包括：*WT1* 与 11p13；*WT2*（OMIM 194071）与 11p15.5；*WT3*（OMIM 194090）与 16q；*WT5*（OMIM 601563）与 7p15-p11.2；*WT4*（OMIM 601363）与 17q12-q21。Wilms 瘤患者的肿瘤组织中有 *WT* 基因的纯合缺失，而正常组织中则为杂合子，其发生机制与视网膜母细胞瘤相同。

（三） 家族性腺瘤性息肉综合征（MIM 175100）

家族性腺瘤性息肉（familial adenomatous polyposis，FAP）是常染色体显性遗传病，以结肠和直肠多发息肉为特征，带有这类家族性腺瘤性息肉病的个体在他们十几岁时便开始向多发性息肉发展，直至产生癌变。典型的家族性腺瘤性息肉患者发展成癌的平均年龄为 39 岁。一些人是此病的突变型，称为衰减型家族性腺瘤性息肉病，此型息肉增长较缓，发生癌变的平均年龄为 55 岁。此外，家族性腺瘤性息肉也可发生于上消化道，肿瘤可转移到脑、甲状腺等身体其他器官。

FAP 典型的临床症状为结肠和直肠的腺瘤性息肉；在肠远端黏膜表面可见很多息肉，这是一种严重的癌前病变，在未经过治疗的致病基因携带者中可见一个或多个息肉发育不良并出现恶性变。肿瘤可发生于儿童期到 70 岁间的任何年龄，其症状为恶病质，如体重减轻、营养不足、肠梗阻及血性腹泻。

位于染色体 5q21-q22 的 *APC* 基因为该病的主要致病基因。*APC* 基因突变可导致典型的和衰减的家族性腺瘤性息肉。这些突变影响细胞的正常生长和功能。由 *APC* 基因突变引起的细胞过度增长可导致结肠息肉。尽管带有 *APC* 基因突变的大多数人最终将发展成结直肠癌，但是息肉的数量及成癌的时间取决于基因突变的部位。

该病曾被命名为多发性结肠息肉、遗传型结肠息肉、家族性多发性息肉及家族性息肉病(familial polyposis coli,FPC)。现根据其病发部位非局限于结肠而称其为家族性腺瘤性息肉病(FAP)。

Gardener 综合征是结肠息肉的另一种类型,具有明显的肠道外症状,其中典型症状为骨瘤、显著的视网膜损伤,现已知该综合征也是由于 *APC* 基因突变所引起的 FAP 的另一种表型。结肠外表现一般可分成三类:发生在上消化道的腺瘤性息肉,是导致患者死亡的主要原因;口腔、皮肤及骨的症状,可作为重要的诊断依据,但通常为良性;消化道外其他器官的恶性改变。

二、常染色体隐性遗传的恶性肿瘤综合征

人类一些以体细胞染色体断裂为主要表现的综合征多具有常染色体隐性遗传特性,统称为染色体不稳定综合征,它们具有不同程度的易患肿瘤的倾向。

(一) Bloom 综合征(MIM 210900)

1. Bloom 综合征的临床特征　又称为 Bloom-Torre-Machacek 综合征,是一种罕见的常染色体隐性疾病,主要特征为患者出现高频率的染色体断裂和重组。本病由皮肤科医师 David Bloom 博士发现并于 1954 年首次报道,患者常见的临床表现包括:身材矮小,慢性感染,免疫功能缺陷,日光敏感性面部红斑和轻度颜面部畸形,面部红疹呈蝴蝶形,独有的面部特征表现为面部窄长、下颌小和耳鼻突出。该病的并发症包括慢性肺部疾病、糖尿病及学习能力差。小部分患者可见智力发育迟缓。最显著的并发症为高度肿瘤易感性且多在 30 岁前发生各种肿瘤和白血病。据美国纽约血液研究中心人类遗传学实验室 Bloom 综合征患者登记处的报道,至 1995 年上半年,已在日本、西班牙、葡萄牙、意大利、希腊、德国、佛兰芒(Flemish),英国、威尔士、荷兰、北美、南美、澳大利亚、非洲、土耳其以及地中海东部诸国家和岛屿(Levantine)的人群中共发现 179 例 Bloom 综合征患者,其中 57 例为犹太籍,从而提示 Bloom 综合征发病具有明显的种族特异性。

2. Bloom 综合征患者的细胞遗传学改变　染色体不稳定性或基因组不稳定性(genomic instability)是 Bloom 综合征患者细胞遗传学的显著特征,主要表现在:①体外培养的 Bloom 综合征细胞株的染色体易发生断裂并易形成结构畸变;体内 Bloom 综合征细胞如颊黏膜细胞在分裂间期常可见细胞内出现多个微核结构;②Bloom 综合征细胞的染色体断裂易发生在染色体的同源序列之间,从而出现频发的姐妹染色单体交换(SCEs)现象;③Bloom 综合征细胞的染色体不但在编码序列之间,而且在非编码序列之间也同样存在断裂性突变;④培养的 Bloom 综合征细胞中常见四射体结构,尤其常见于短期培养的 Bloom 综合征淋巴细胞中,但在正常人的细胞中却罕见。

3. Bloom 综合征的其他特征　Bloom 综合征细胞中出现大量异常细胞特性改变,如细胞中尿嘧啶 DNA 糖基酶表现等电点改变,耐热差,K_m、V_m 改变,同时细胞对 5-氟尿嘧啶、5-溴尿嘧啶及嘌呤类似物等不敏感,尤其是只有嘧啶类似物达到 10 ~100 倍浓度才能抑制其活性。在紫外线、X 线或限制酶损伤 DNA 后,正常哺乳动物细胞核中会有 P53 蛋白堆积,而 Bloom 综合征细胞在紫外线照射和 SV40 感染后无 P53 聚集,X 线照射后 P53 堆积延缓。Bloom 综合征患者缺乏美洲商陆有丝分裂原(pokeweed mitogen,PWM),导致淋巴细胞激活途径改变。此外,Bloom 综合征细胞中 DNA 连接酶存在异常,DNA 连接酶 Ⅰ 基因突变可能是 Bloom 综合征代谢缺陷的重要原因。Bloom 综合征细胞主要生化缺陷表现为产生过量的过氧化物阴离子,其结果导致姐妹染色单体交换异常和染色体损伤。

4. Bloom 综合征基因定位　German 依据 27 例 Bloom 综合征患者家系调查证实 Bloom 综合征为一种常染色体隐性遗传病。但直到 1992 年 McDanel 等才利用微细胞介导的染色体转移技术(microcell-mediated chromosome transfer,MMCT)把 Bloom 综合征的编码基因(*BLM* 基因)定位在 15 号染色体长臂。而 German 等利用纯合性定位技术将 *BLM* 基因限定在 15q26.1 区带内。Ellis 等克

隆了 *BLM* 基因的全长 cDNA，且发现 *BLM* 基因的突变是 Bloom 综合征患者发病的分子遗传学基础。

5. Bloom 综合征的分子遗传学诊断　目前通用的分析技术是利用 Bloom 综合征患者的体外培养细胞获得 *BLM* 基因的 cDNA，再用相关引物扩增 *BLM* 基因的全长 cDNA，SSCP 技术筛查所获得的 PCR 扩增片段，对出现 SSCP 改变的 PCR 片段加以序列分析，从而鉴定 *BLM* 基因编码序列上存在的各种突变。

6. Bloom 综合征的细胞遗传学诊断　由于体外培养的 Bloom 综合征患者的细胞姐妹染色单体交换发生率显著增高，因此显著增高的 SCEs 就成为 Bloom 综合征患者确诊不可缺少的细胞遗传学依据。正常细胞在含有 5-溴尿嘧啶的培养基中单一细胞中期分裂象 <10，但 Bloom 综合征患者单一细胞中期分裂象介于 50 ~ 100 之间，但 SCEs 的大小可因选择的细胞类型或不同的 Bloom 综合征患者而存在一定的差异。常用的细胞类型如下：短期培养的 T 淋巴细胞、B 淋巴细胞和骨髓细胞；EB 病毒转化的 B 淋巴细胞，长期培养的 T 淋巴细胞和各种组织来源的成纤维细胞，包括 SV40 转化的细胞系。

（二）Fanconi 贫血（MIM 227650）

Fanconi 贫血（Fanconi anemia，FA）在临床上相当罕见，1927 年由瑞士科学家 Fanconi 首先对该病进行描述，它属于常染色体隐性遗传病，发病率大约为 1/350 000。主要表现为各类起源于骨髓干细胞的血细胞发育受阻（全血细胞减少症），最终导致贫血、白血病及血小板减少症。贫血多发生在 5 ~ 10 岁，多数病例伴有先天性畸形，特别是骨骼系统，如拇指短小或缺如、多指、桡骨缩短、体格矮小、小头、眼裂小，少数伴有肾畸形及心血管畸形等，皮肤色素沉着及咖啡牛奶样色斑也很常见。儿童期肿瘤发生危险性也增高，特别是白血病的发生率显著增高，尤其易患急性单核细胞性白血病。

在培养的 FA 细胞中普遍存在染色体不稳定，染色体自发断裂率明显增高，单体断裂、裂隙等染色单体畸变很多，双着丝粒体、片段、核内复制也很常见。研究发现，FA 患者的染色体畸变类型要多于 Bloom 综合征，并且 Bloom 综合征细胞染色体的交叉互换多发生在同源染色体之间，而 FA 则多发生在非同源染色体之间。另外，通过 FISH 研究发现，FA 细胞在端粒序列-TTAGGG 处多发生断裂，因此，端粒完整性的破坏及其保护功能的丧失也增加了染色体的不稳定性。分子遗传学研究发现，FA 细胞的 DNA 错配修复系统中可能存在核酸外切酶或 DNA 连接酶活性的缺陷，也有人发现 FA 细胞质中存在高水平的 DNA 拓扑异构酶 I，而它在细胞核中的含量却很低，因此推测可能是 DNA 修复相关酶类在由胞质到胞核的运输通路上存在缺陷。无论是哪种细胞缺陷，都使 FA 患者染色体的不稳定性增高、对 DNA 损伤敏感性增强、等位基因易发生缺失突变等，损害了基因组的完整性，最终导致对白血病等恶性肿瘤的易感性增强。

FA 细胞主要对一些化合物如丝裂霉素 C、双环氧丁烷、顺铂反应敏感，并由此导致在一条 DNA 链内或两条互补链之间的两个核苷酸之间形成交联。一般情况下，FA 细胞在一定浓度的交联剂存在时停止生长并死亡，而野生型细胞则很少出现这种情况。检测培养细胞对丝裂霉素 C 的敏感性是诊断 FA 的一个有效方法，也有人尝试利用 DEB 诱导染色体断裂对 FA 进行产前诊断。

目前发现有 13 个已知的基因突变会导致 FA：*FANCA*、*FANCB*、*FANCC*、*FANCD1*、*FANCD2*、*FANCE*、*FANCF*、*FANCG*、*FANCI*、*FANCJ*、*FANCL*、*FANCM* 和 *FANCN*。*FANCB* 是 FA 一个例外的致病基因，因为 FA 是常染色体隐性遗传病，但 *FANCB* 定位于 X 染色体上。

目前，全世界大约有 1 000 人患有此病。回顾性研究发现，FA 的转归很差，多数将导致白血病及实体瘤的发生，如阴道癌、食管癌及头颈部肿瘤。

（三）共济失调毛细血管扩张症（MIM 208900）

共济失调毛细血管扩张症（ataxia telangiectasia，AT）是一种罕见的常染色体隐性遗传病，发病率为1/100 000～1/40 000。其主要的临床表现是进行性小脑共济失调，肺部反复感染，以及眼和面部皮肤的毛细血管扩张。其他特征包括对射线的杀伤作用异常敏感、染色体不稳定性增加、易患肿瘤和免疫缺陷等。以前人们一直认为该病是多基因遗传病。但经过近40年的研究，AT的致病基因终于在1995年被克隆，同时也确定AT为单基因遗传病。此后，在*AT*基因结构组成、功能及与肿瘤发生的相关性研究等方面也取得了一定的进展。

1995年，Savitsky报道，在其研究的AT患者中，均有一个相同的基因发生突变，说明该基因即为AT的致病基因，也说明AT为单基因遗传病。Savitsky将该基因命名为*AT*基因，并克隆了该基因的全部序列。*AT*基因是迄今为止发现的外显子最多的人类基因之一，也是最重要的基因之一。*AT*基因全长约150kb，编码序列约12kb，共有66个外显子，编码一个有3056个氨基酸残基、分子量为350kD的蛋白质。

自*AT*基因克隆以来，对其功能的研究取得了一定的进展。AT是一种多系统综合征，累及多种组织和细胞，如AT患者的小脑Purkinje细胞、成纤维细胞及淋巴细胞等都表现出对射线的异常敏感，因此*AT*基因被视为管家基因（housekeeping gene）。就目前所知，野生型*AT*基因具有下列功能：①与DNA损伤修复有关；②抑制细胞凋亡；③控制免疫细胞对抗原的反应；④介导细胞对胰岛素的反应；⑤阻止基因重排（gene rearrangement）；⑥可能与性成熟有关。根据Savitsky等（1996）的报道，*AT*基因产物羧基端的400余个氨基酸与PI3K（磷脂酰肌醇3-激酶）的催化功能域有高度的序列同源性。PI3蛋白家族包括多种蛋白质，如TELL、MEC1、RAD3、MEI41、DNA-PKCS及TOR等。这些蛋白质参与调控细胞周期、控制端粒酶长度和应答DNA损伤。将*AT*基因产物的序列与这些蛋白质进行比较，可以推测其功能。

到目前为止，在AT患者中已发现了30多种突变，这些突变分布于*AT*基因的整个编码序列，其中绝大多数突变会造成*AT*基因的截短或大片段缺失，从而导致AT蛋白失活。AT患者易患多种肿瘤，肿瘤发病率比正常人群高近10倍，其中尤以淋巴瘤和白血病最为常见。

（四）着色性干皮病（MIM 278700）

着色性干皮病（xeroderma pigmentosum，XP）是一种罕见的、致死性常染色体隐性遗传病，发病率为1/250 000。XP的主要临床特点为早发的起源于皮肤上皮外层细胞（鳞状细胞）或内层细胞（基底细胞）的皮肤癌，XP患者也易患其他类型肿瘤包括恶性黑色素瘤、角化棘皮瘤、肉瘤、腺癌。此外，有些患者还伴有生长迟缓、性发育不良、智力障碍、小头和神经性耳聋等神经方面的表现。XP患者对光极敏感，皮肤、眼和舌部易受损，皮肤有许多色素斑点，也常常是皮肤癌的发生部位。XP患者最常见的死因是转移性的恶性黑色素瘤和鳞状细胞癌，患者很少能活过20岁。

紫外线辐射促使DNA相邻嘧啶形成稳定的共价连接，如T-T，C-T，C-C等，在这些嘧啶二聚体中，T-T出现频率最高。除此之外，一些化合物也具有核苷酸交联剂的作用，另一些化合物可为DNA碱基添加化学基团。二聚体核苷酸交联和特异的核苷酸侧基破坏了染色体结构并导致其突变。正常情况下，核苷酸切除修复（NER）系统可切除这些受损的核苷酸并重建正常核苷酸序列。例如经紫外线辐射后，NER系统蛋白质结合到DNA上胸腺嘧啶二聚体，NER系统的DNA解旋酶在二聚体处向相反两个方向解开DNA双螺旋，核酸内切酶切断二聚体及两侧的受损DNA单链，包括清除二聚体的寡核苷酸片段，然后由DNA聚合酶填补这段缺口，再由DNA连接酶封闭这条DNA链。

由于XP患者核苷酸切除修复途径缺陷，细胞对紫外线辐射高度敏感。被紫外线照射之后，XP患者皮肤发红，随之发生萎缩和毛细血管扩张，这些部位逐渐变成疣状，最终发生多发性皮肤恶性

肿瘤。

真核细胞中NER系统具有30多种蛋白质，不同XP患者的临床差异提示该病存在遗传异质性。用不同患者的成纤维细胞进行的细胞杂交确定了XP患者至少分为8种互补型，称为*XPA*至*XPG*以及*XPV*，推测至少有8种不同的基因产物参与NER。绝大多数已知的*XP*基因突变使NER系统失去功能并导致紫外线诱导的细胞癌变。

Summary

Recognized as a neoplasm formed by a group of cells with uncontrolled proliferating ability, tumors can be categorized into benign tumors and malignant tumors. A malignant tumor is a genetic disease that occurs within somatic cells and is triggered by structural and formational alterations of tumor-related genes(including oncogenes and tumor suppressor genes). These genes are normally implicated in the regulation of cell growth and the repair of gene mismatch. On the other hand,5% of cancer is caused by the presence of a germline mutation and can be passed on to the next generation.

Despite the scientific expansion of the 21st century, our understanding of tumor etiology remains incomplete. Furthermore, the cancer mortality increases significantly, even with good control of infectious diseases and emergence of longevity. This increase in mortality is also caused by environmental pollution, stress and unhealthy life habits such as smoking. Regardless, no apparent determinant factor has been identified so far in the pathogenesis of most malignant tumors, in which genetic basis is believed to play a critical role. It would be relatively easy to understand the initiation and promotion of cancer and therefore design a therapeutic strategy if only cancer was a disease involving a single gene mutation. Unfortunately, this is not the case; any misstep in the induction of proliferation, the regulation of cell cycle and/or programmed cell death(apoptosis) will inevitably lead to the development of a tumor.

（傅松滨）

参考文献

1. http://www.ncbi.nlm.nih.gov:Online Mendelian Inheritance In Man
2. http://en.wikipedia.org/
3. Burkhart DL, Sage J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. Nat Rev Cancer, 2008, 8(9): 671-682
4. Becker-Catania SG, Gatti RA. Ataxia-telangiectasia. Adv Exp Med Biol, 2001, 495: 191-198
5. Dworaczek H, Xiao W. Xeroderma pigmentosum: a glimpse into nucleotide excision repair, genetic instability, and cancer. Crit Rev Oncog, 2007, 13(2): 159-177
6. Fred Bunz. Principles of Cancer Genetics. Springer, 2008
7. Gratchev A. The nucleotide excision repair of DNA in human cells and its association with xeroderma pigmentosum. Adv Exp Med Biol, 2008, 637: 113-119
8. Halazonetis TD, Gorgoulis VG, Bartek J. An oncogene-induced DNA damage model for cancer development. Science, 2008, 319(5868): 1352-1355
9. Jeter JM, Kohlmann W, Gruber SB. Genetics of colorectal cancer. Oncology(Williston Park), 2006, 20(3): 269-276
10. Macrae F, du Sart D, Nasioulas S. Familial adenomatous polyposis. Best Pract Res Clin Gastroenterol, 2009, 23(2):

197-207

11. Manning AT, Garvin JT, Shahbazi RI, Miller N, McNeill RE, Kerin MJ. Molecular profiling techniques and bioinformatics in cancer research. Eur J Surg Oncol. ,2007,33(3):255-265
12. Malik K, Yan P, Huang TH, Brown KW. Wilms' tumor: a paradigm for the new genetics. Oncol Res. ,2001,12:441-449
13. Poulaki V, Mukai S. Retinoblastoma: genetics and pathology. Int Ophthalmol Clin, 2009, 49(1):155-164
14. Pasche B. Recent advances in breast cancer genetics. Cancer Treat Res, 2008, 141:1-10
15. Sverre Heim, Felix Mitelman. Cancer Cytogenetics. 3rd ed. Wiley-Blackwell, 2009
16. Thomas D G. Principles of Medical Genetics -second Edition. Williams&Wilkins-a waverly company, 1999
17. Wolfgang A S. Molecular Biology of Human Cancers (An Advanced Student's Textbook). Springer, 2007
18. Zender L, Lowe SW. Integrative oncogenomic approaches for accelerated cancer-gene discovery. Curr Opin Oncol, 2008, 20(1):72-76

第十二章　临床遗传学

临床遗传学(clinical genetics)是医学遗传学的重要组成部分,是医学遗传与遗传医学(genetic medicine)相交叉的领域。其内容包括遗传病的诊断、治疗和预防。

第一节　遗传病的诊断

遗传病的诊断是项复杂的工作,它需要各个学科的密切配合。遗传病的诊断包括常规诊断和特殊诊断。常规诊断指与一般疾病相同的诊断方法,特殊诊断指利用遗传学的方法,如染色体和染色质检查、家系分析等方法进行诊断。而且,遗传病的特殊诊断往往是确诊的关键。目前,临床上的遗传病诊断主要包括:临症诊断(symptomatic diagnosis)、症状前诊断(presymptomatic diagnosis)和出生前诊断(prenatal diagnosis)。

一、临症诊断

临症诊断是医务工作者根据已出现症状的患者的各种临床表现进行分析,并进行疾病的诊断和遗传方式的判断,是遗传病临床诊断的主要内容。

(一) 病史、症状和体征

1. 病史　遗传病多有家族聚集现象,由此病史的采集极为重要,采集过程中要遵循准确、详细的原则。另外,还要根据不同的遗传病进行特别的调查。

2. 症状和体征　遗传病有和其他疾病相同的症状和体征,往往又有其本身特异性症候群,为诊断提供线索。由于大多数遗传病在婴儿或儿童期即可有体征和症状表现,故除观察外貌特征外,还要

注意身体发育快慢、智力增进情况、性器官及第二性征发育是否异常等。

（二）系谱分析

在遗传病诊断时进行系谱分析有助于区分单基因病与多基因病，以及属于哪种遗传方式，有助于区分某些表型相似的遗传病以及由于遗传异质性而出现的不同遗传方式。系谱分析时应注意：系谱的系统性、完整性和可靠性；分析显性遗传病时，应注意对已有延迟显性的年轻患者，由于外显不全呈隔代遗传时，不要误认为是隐性遗传；有些遗传病家系除先证者外，家庭成员中找不到其他患者，此时应考虑是否为新的基因突变；要注意显性与隐性概念的相对性，同一遗传病可因观察指标不同而得出不同的遗传方式，从而导致发病风险的错误估计。

（三）细胞遗传学检查

细胞遗传学检查，即染色体检查或核型分析，是较早应用于遗传病诊断的辅助手段。目前随着显带技术的应用以及高分辨染色体显带技术的出现和改进，能更准确地判断和发现更多的染色体数目和结构异常综合征，还可以发现新的微畸变综合征，因而该方法是确诊染色体病的主要方法。但是在临床应用上，进行染色体分析的指征很难掌握，因此异常核型的检出率常常较低，造成人力和物力的浪费，但控制过严可能使一些阳性病例漏诊。

染色体检查标本的来源，主要取自外周血、绒毛、羊水中脱落细胞和脐血、皮肤等各种组织。

染色体检查的指征：有明显的智力发育不全；生长迟缓或伴有其他先天畸形者；夫妇之一有染色体异常，如平衡结构重排、嵌合体等；家族中已有染色体异常或先天畸形的个体；多发性流产妇女及其丈夫；原发性闭经和女性不育症；无精子症男子和男性不育症；两性内外生殖器畸形者；疑为先天愚型的患儿及其父母；原因不明的智力低下伴有大耳、大睾丸和多动症者；35岁以上的高龄孕妇。

1. 染色体显带技术　20世纪70年代前后发展起来的染色体显带技术是细胞遗传学中的一大突破。最基本的是G显带，即制备中期染色体标本，用胰酶处理后以吉姆萨溶液染色，最后在光学显微镜下观察深浅相间的区和带，该技术称G显带。根据对染色体处理方法和染料的不同，先后又发展了10余种显带技术，包括：Q显带（氮芥喹吖因等染色，带型与G带相同），R显带（用加热、荧光或其他处理获得与G带深浅相反的带），T显带（显示端粒），C显带（显示着丝粒），N显带（显示核仁组织区）以及最新的限制性内切酶显带。

利用染色体显带技术，可以使许多疾病在染色体水平找到原发性改变，如肿瘤、发育缺陷、心血管疾病等。通过显带技术，可以把与此疾病的相关基因确定在一个较小的范围内，以利进一步研究。

2. 染色体原位杂交（chromosome in situ hybridization）　应用标记的DNA片段（标记物可为生物素、地高辛、荧光等）与玻片上的细胞、染色体或间期核的DNA或RNA杂交，在这些核酸不改变原来结构的情况下，研究核酸片段的位置、相互关系，因此称为原位杂交。用生物素、地高辛等标记物标记的DNA探针进行原位杂交后，用荧光染料（喹吖因、罗丹明、FITC等）标记的生物素亲和蛋白和抗亲和蛋白的抗体进行免疫检测和放大，使探针杂交的区域发出荧光，这种原位杂交称为荧光原位杂交（FISH）。FISH技术具有快速、经济、安全、灵敏度高、特异性强等优点。它自问世以来，已广泛应用于细胞遗传学、基因定位和基因制图等领域中。

（四）生化检查

生化检查是遗传病诊断中的重要辅助手段，包括一般的临床生化检验和遗传病的特异检查。不同类型的遗传病的缺陷不同，因此生化检查也各种各样，用于分析酶变型的方法主要有电泳速率、酶动力学、指纹分析和免疫反应等常用技术；用于分析蛋白质变型的方法主要靠电泳技术、肽链和氨基酸顺序分析来辨认。另外，测定中间代谢产物也有助于诊断代谢病，比如，通过测定尿中苯丙酮酸或苯乙酸可诊断苯丙酮尿症。

遗传性代谢病是一组先天性生化紊乱所致的疾病，属单基因病。目前已知的200余种遗传性代

谢病中，多数为酶缺陷病，少数为非酶缺陷病。酶缺陷病中大部分为常染色体隐性遗传，个别为X连锁隐性遗传。各种遗传性代谢病的发生率都很低，一般在1/100 000～1/50 000，它在不同地区、不同种族中发病率的差异很大。遗传性代谢病是基因突变的结果，基因表达调控异常、基因缺失、基因点突变导致蛋白质缺如和结构异常或翻译后加工修饰缺陷等，即可导致酶蛋白的缺如或功能异常。目前临床上常用的生物化学检查方法适用于检测酶的缺陷和代谢中间产物。血和尿液由于易采集，及其方法学的不断改进，一直受到检测者的采用，目前已制成滤纸片和显色反应进行检测。

（五）基因诊断

1. 限制性片段长度多态性（restriction fragment length polymorphism，RFLP） 人群中不同个体基因的核苷酸序列存在差异，称为DNA多态性。DNA顺序上发生变化而出现或丢失某一限制性内切酶位点，使酶切产生的片段长度和数量发生变化称为RFLP。任何一个基因内大片段的缺失、插入以及基因重排，即使不影响到限制性内切酶位点的丢失或获得，也很可能引起限制性内切酶图谱的变化，使限制性酶切片段的大小和数量发生变化，因而这类基因突变可以通过限制性内切酶DNA或结合基因探针的杂交的方法将突变找出。例如镰形红细胞贫血症的基因诊断（图12-1）、血友病A的诊断（图12-2）。

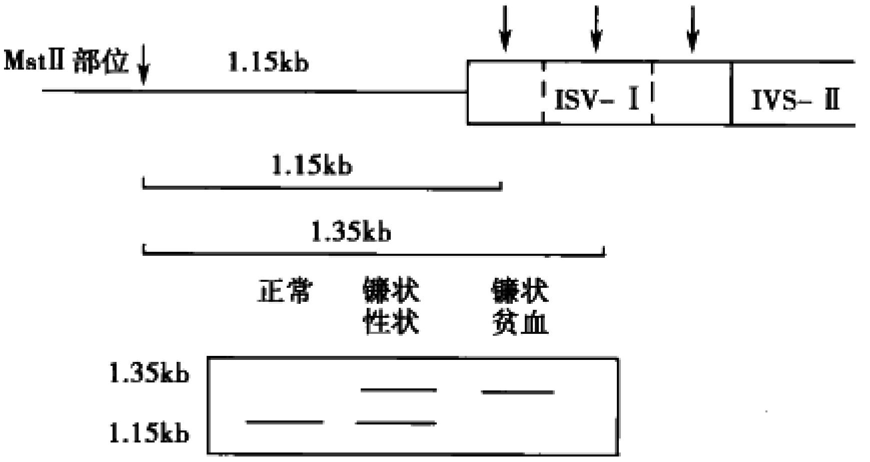


图12-1 镰形红细胞贫血症的基因诊断

已知镰形红细胞贫血症的突变基因是编码β珠蛋白链的第6位密码子由GAG变为GTG，可用限制性内切酶MstⅡ进行检测。因为这一突变使正常存在的MstⅡ切点消失，这就使正常情况下存在的1.15kb及0.2kb条带变成患者（纯合子）的1.35kb条带（图12-1）。

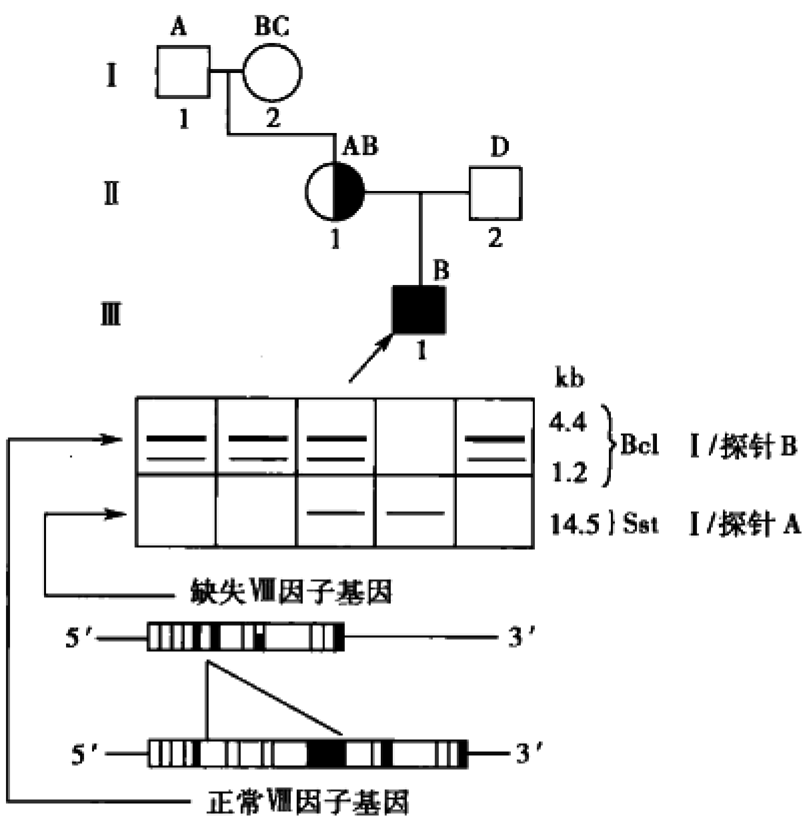


图12-2 Ⅷ因子基因部分缺失所致血友病A的基因诊断
探针A是Ⅷ因子基因1.7kb KpnⅠcDNA片段，包括1～12外显子。探针B是Ⅷ因子基因4.7kb EooRⅠcDNA片段，包括14～25和部分26外显子。A、B、C、D代表X染色体。患者无Bcl/探针1.2kb和4.4kb，但有SstⅠ/探针A14.5kb

血友病A是一种X连锁隐性遗传病。用Ⅷ因子基因的cDNA片段作为探针对待检者DNA酶切片段进行杂交，就可检出Ⅷ因子基因部分缺失的男性患者和女性携带者。图12-2家系Ⅲ-1患有严重的血友病A，经Ⅷ因子治疗后产生Ⅷ因子抑制物。用两种不同的内切酶和探针组合对家系成员作RFLP分析。应用BclⅠ/探针B（Ⅷ因子基因4.7kb EcoRI cDNA片段，包括14～25和部分26外显子），患者不见1.2kb和4.4kb。应用SstⅠ/探针A（Ⅷ因子基因1.7kb KpnⅠ cDNA片段，包括1～12外显子），患者和他的母亲（必然杂合子）出现异常的14.5kb。这一实验结果表明，Ⅷ因子基因3′端部分缺失后，其5′端残留部分可由SstⅠ/探针A检出。杂合子兼有BclⅠ/探针B检出的1.2kb和4.4kb与SstⅠ/探针A检出的14.5kb。

2. PCR（聚合酶链反应）及相关技术

（1）聚合酶链反应（polymerase chain re-

action，PCR)：PCR 通过变性、退火、延伸的循环周期，使特定的基因或 DNA 片段在短短的 2～3 小时内扩增数十万至百万倍，大大缩短了诊断时间。PCR 常结合其他技术进行诊断。

(2) PCR 相关技术：以 PCR 为基础的相关技术有多种，巢式 PCR(nest PCR)、增效 PCR(booster PCR)、多重 PCR 等，本节着重介绍 PCR/ASO、PCR-SSCP、PCR-DGGE 技术。

1) PCR/ASO 探针斑点杂交：等位基因特异的寡核苷酸探针杂交是最早用来检测点突变的方法。用人工合成的 19 个碱基左右长度的 ASO 探针，在严格的杂交洗脱温度下，可区分一个碱基的差别，用针对正常和突变的 ASO 可准确鉴定个体的基因型。使用 TMA 杂交系统，洗脱条件只与 ASO 的长度有关，而与 ASO 碱基组分无关，这样可将同一基因的不同突变的 ASO 探针固定在滤膜上反向杂交，通过标记待测个体 DNA 的 PCR 扩增产物，即可同时鉴定待测个体的基因型。

2) PCR-单链构象多态性(PCR-single strand conformation polymorphisms，PCR-SSCP)：DNA 单链构象多态性是指等长的单链 DNA 因核苷酸序列的差别而产生构象差异，在非变性聚丙烯酰胺凝胶中表现为电泳迁移率的差别。将突变所在区域的 DNA 片段进行 PCR 扩增后进行电泳，根据单链条带位置的改变判断某个体是否存在特异的突变。以 SSCP 为线索还可通过扩增片段的直接测序确定突变位点。但并不是所有的核苷酸序列改变都引起单链构象改变，因此 SSCP 并不能鉴别所有突变。

3) PCR 产物变性梯度凝胶电泳(denaturing gradient gel electrophoresis，DGGE)分析：DGGE 通常要用尿素甲酰胺浓度梯度，电泳在衡温水浴(＞30℃)中进行。正常等位基因与突变等位基因因核苷酸不同而解链温度(即解离强度)不同，电泳过程中 DNA 逐渐由低强度进入高强度，解离温度较低的 DNA 片段先达到其解离强度(变性梯度)而解离成局部单链，泳动减慢；解离温度高的 DNA 片段保持原来的泳速继续前进，直到达到解离强度时才发生解离，速度减慢，通过溴化乙啶或银染观测区带的变化。

3. DNA 测序　DNA 序列测定方法的诞生为详细分析遗传病等疾病的基因结构与功能奠定了基础。目前 DNA 序列测定自动化已实现了分析反应自动化和读片自动化。使用 4 种不同颜色的荧光染料的无放射性标记的 DNA 测序法及毛细管电泳的方法，是 DNA 序列分析自动化研究方面的一个重要进展。

利用 DNA 测序技术可用来检测基因确定的突变部位与类型，是目前最根本的一种检测基因突变的方法。如检测基因片段的缺失或插入、动态突变(三核苷酸重复序列的扩增所致)等。

4. DNA 芯片　DNA 芯片(DNA chip)又称微阵(microarray)，属于生物芯片的一种。它是把上万种寡核苷酸或 DNA 样品密集排列在玻片、硅片或尼龙膜等固相支持物上，通过激光共聚焦荧光显微镜获取信息，电脑系统分析处理所得资料，一次微排列可对上千种甚至更多基因的表达水平、突变和多态性进行快速、准确的检测。

DNA 芯片技术是一种高效准确的 DNA 序列分析技术。基于 PCR 技术的检测方法(如 PCR-SSCP、PCR-DGGE 等)大多用于检测突变是否存在，而不能确定突变性质。DNA 测序法能够准确确定突变的部位与性质，但目前的以凝胶电泳为基础的测序技术费时费钱，所以在实际应用中较少采用直接测序来检测突变。DNA 芯片应用于检测基因突变，不仅可以准确地确定突变位点和类型，它的快速高效是目前的直接测序所无法比拟的，它可以同时检测多个基因乃至整个基因组的所有突变。

CheeM 等合成了含有 135 000 个长度为 25bp 的探针的 DNA 芯片，并用它分析了 16.6kb 的人类线粒体基因组(mt DNA)。共分析了 10 个样本，检测出 505 个多态位点，并且在 Leber 遗传性视神经病患者的 mt DNA 中成功地检测出了 3 个致病性突变位点。

Cronin 等分别用两种 DNA 芯片检测囊性纤维化跨膜传导调节基因(cystic fibrosis transmembrane conductance regulator，CFTR)的突变。其中一个芯片检测了 CFTR 基因的第 10 个和 11 个外显子的已知突变，包括缺失、插入和单碱基转换。并分析了 10 个未知样品的 CFTR 基因，其结果与 PCR-RFLP 的分析结果完全一致。DNA 芯片还被用来检测 β 地中海贫血患者的基因突变，并在 β-珠蛋白基因上

检测出了3个突变位点。

DNA芯片除了用于检测基因突变外，还可用于遗传作图，杂交测序以及监测基因的表达水平等。

5. Western印迹技术 Western印迹技术是检测特定的蛋白质的方法。X连锁隐性遗传的Duchenne肌营养不良(DMD)是由于基因缺陷而使肌细胞增强蛋白(dystrophin)合成异常，用Western法可检测患者肌细胞增强蛋白，从而对DMD进行诊断。

二、症状前诊断

某些常染色体显性遗传病的杂合子个体往往发病年龄延迟，比如Huntington舞蹈病杂合子的好发年龄在40岁左右，而这时的杂合子个体已经生儿育女，他们有1/2的机会将致病基因传给子代，造成子代得病。如能在可疑杂合子个体生育之前就做出诊断，就能避免影响子代。

常染色体显性杂合子个体的症状前诊断主要依赖于家系调查和系谱分析，依赖于各种临床检查和实验室检查，也依赖于DNA诊断技术的应用。通过家系调查和系谱分析，可估计出家系中各成员的杂合子风险。对风险较高的个体，应作进一步检查，以明确诊断。目前，在症状出现前能明确诊断的方法只有DNA检查。

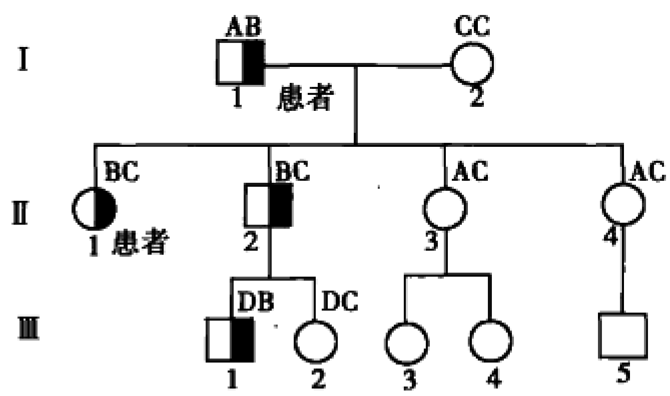


图12-3 Huntington舞蹈病家系成员症状前连锁DNA分子单倍型诊断

该家系Huntington舞蹈病与分子单倍型B(Hind Ⅲ/G8探针：4.9kb、17.5kb)连锁，Ⅱ-2和Ⅲ-1为症状前患者

下面以Huntington舞蹈病为例说明DNA分析在常染色体显性杂合子症状前诊断中的应用。图12-3示一个Huntington舞蹈病家系。图中Ⅰ-1和Ⅱ-1已经发病，其他成员尚未得病，他们是否也带有致病基因呢？DNA分析可作出诊断。患者和家系部分成员的DNA以HindⅢ酶切，用Huntington舞蹈病连锁G8探针进行分子杂交，检出A、B、C、D四种分子单倍型，如表12-1所示。患者Ⅰ-1和Ⅱ-1共有分子单倍型B(4.9kb、17.5kb)，提示分子单倍型B与Huntington舞蹈病基因连锁。分析家系中其他成员，Ⅱ-2和Ⅲ-1也含有分子单倍型B，在不发生重组的情况下，Ⅱ-2和Ⅲ-1为杂合子，他们迟早要发病，应尽快作好预防工作。

表12-1 Huntington舞蹈病连锁G8探针分子单倍型

分子单倍型	Hind Ⅲ酶切部位1	酶切片断(kb)	Hind Ⅲ酶切部位2	酶切片断(kb)
A	−	17.5	+	3.7
B	−	17.5	−	4.9
C	+	15.0	+	3.7
D	+	15.0	−	4.9

三、出生前诊断

出生前诊断或产前诊断是以羊膜穿刺术和绒毛膜取样等技术为主要手段，对羊水、羊水细胞及绒毛膜进行遗传学分析，以判断胎儿的染色体或基因等是否正常。如果确认正常胎儿则继续妊娠至足月生产，是预防遗传病患儿出生的有效手段。

(一) 出生前诊断的对象

根据遗传性疾病的严重程度和发病率的高低，可将出生前诊断的对象排列如下：①夫妇之一有染色体畸变，特别是平衡易位携带者，或夫妇核型正常，但曾生育过染色体病患儿的夫妇；②35岁以上

的高龄孕妇；③夫妇之一有开放性神经管畸形，或是生育过这种畸形儿的孕妇；④夫妇之一有先天性代谢缺陷，或生育过这种患儿的孕妇；⑤X 连锁遗传病基因携带者孕妇；⑥有原因不明的习惯性流产史的孕妇；⑦羊水过多的孕妇；⑧夫妇之一有致畸因素接触史的孕妇；⑨具有遗传病家族史，又系近亲婚配的孕妇。

（二）出生前诊断的方法与应用

出生前诊断的常用方法包括非侵袭性和侵袭性方法，下面只着重介绍几种常用的方法。

1. 非侵袭性方法　它包括母亲血清与尿液分离、B 超检查、X 线、CT 及磁共振等。B 超与 X 线检查属于影像学检查，它们是出生前诊断的重要手段。

（1）B 超：能详细地检查胎儿的外部形态和内部结构，可通过某些细胞微改变提示染色体异常，使许多胎儿的遗传性疾病得以早期诊断。B 超可进行如下诊断：中枢神经系统异常主要包括神经管缺陷（NTD）、脑积水、小脑畸形等；面、颈部异常，如唇、腭裂和颈部囊状淋巴管瘤等；先天性心脏病；胸部异常包括支气管、肺发育畸形，先天性膈疝，膈膨出和胸腔积液等；染色体异常，已有报道证实超声检查的某些征象与染色体异常有关，如股骨短小和颈部皮褶增厚提示 21 三体征，其敏感性为 82%，特异性为 98%；脐动脉血流异常或单根脐动脉均提示染色体异常；肢体缺陷；其他如先天性肾缺如、肾囊肿、先天性巨结肠等。由于 B 超对胎儿和孕妇基本无损害，因此 B 超检查为目前首选的诊断方法。

（2）X 线及其他：胎儿骨骼在妊娠 20 周后开始骨化，所以在妊娠 24 周后对胎儿进行 X 线检查，最为适宜。诊断剂量的 X 线照射，对胎儿并无不良影响。X 线摄片检查可诊断无脑儿、脑积水、脊柱裂等骨骼畸形。

2. 侵袭性方法　主要包括羊膜穿刺法、绒毛取样法、脐带穿刺术、胎儿镜检查、经皮取血等。不同怀孕阶段采用不同的取样方法，一般孕早期取绒毛，孕中期经羊膜腔穿刺取羊水、经胎儿镜取胎儿标本、直接经腹壁取脐静脉血等。

（1）羊膜穿刺法（amniocentesis）：羊膜穿刺技术（图 12-4）是指在 B 超监视下，用消毒注射器取胎儿羊水的方法，它是产前诊断最基本的方法之一。该法适用于诊断染色体病、遗传性代谢病、神经管缺陷（NTD）和遗传病的 DNA 检测。羊膜穿刺一般在妊娠 16～20 周时进行。可对抽取的羊水进行性别测定、染色质检查和核型分析、DNA 分析等。羊膜腔穿刺的操作是比较安全的，胎儿丢失的风险率很低（约 0.5%～1%）；发生感染和血肿则更罕见；其他妇科并发症则更少。

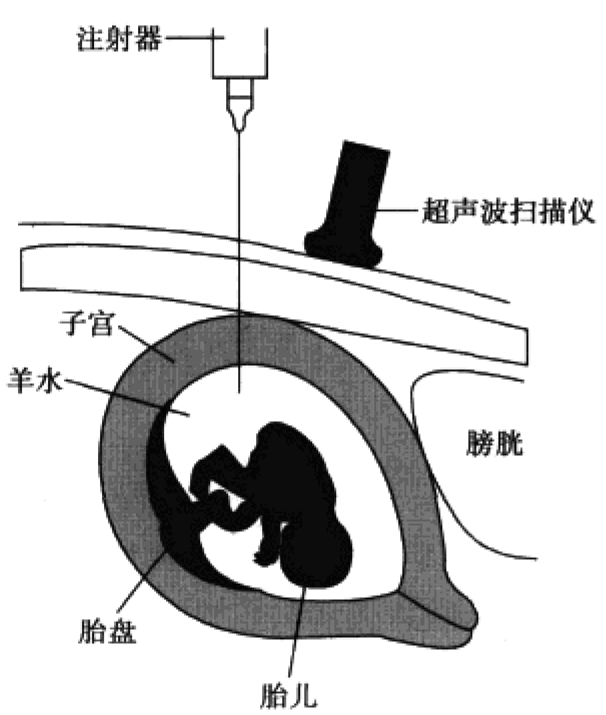


图 12-4　羊膜腔穿刺术

（2）绒毛取样法（chorionic villus sampling，CVS）：绒毛取样技术在妊娠早期诊断中最常见，一般在妊娠 7～9 周时进行。它是指在 B 超的监视下，用一特制的塑料或金属导管从阴道经宫颈进入子宫，再沿子宫壁到达预定的取样位置，并用内管吸取绒毛。但经宫颈取样有易致标本污染、胎儿或母体感染以及操作不便等缺点；也有人采用经腹壁获取绒毛的方法，因为该途径感染的风险低。获得绒毛后可作胎儿性别鉴定、核型分析、生化检查和 DNA 分析。绒毛取样法的优点是检查时间早，需要作出选择性流产时，不会给孕妇带来更多的损伤和痛苦。

（3）脐带穿刺术（cordocentesis）：是指 B 超监视下，用一细针经腹壁进入胎儿脐带并抽取胎儿血样。取样最好在妊娠 18 周。该方法引起流产的概率大约为 1%，低于羊膜穿刺（2.5%）和绒毛取样（7%）。

（4）胎儿镜检查（fetoscopy）：又叫羊膜腔镜或宫腔镜检查，它可在进入羊膜腔后直观观察胎儿的外形、性别、有无畸形等，又可抽取羊水或胎血做各种检

查，还可进行宫内治疗。因此，理论上这是一种最理想的方法。然而由于操作困难，并易引起多种并发症，还不易被医护人员所接受。胎儿镜的最佳取样时间是 18～20 周。

（5）新进展：遗传病的诊断技术已经发展到在受精 6 天胚胎着床之前即可进行。这些技术使孕妇既可避免分娩患遗传病的婴儿，又不必进行流产术，在器官形成之前就对受累胎儿加以识别并进行早期基因治疗。但目前过高的费用和技术本身的需要限制了一些很专业的研究中心对这些技术的应用。

总之，临床医师可以根据妊娠时间及检测目的有针对性地选用不同的穿刺体内取样技术。一般来讲，羊膜穿刺或绒毛取样可用来诊断染色体疾病、遗传代谢疾病、胎儿性别鉴定、所有可用 DNA 方法检测的疾病和开放性神经管缺陷（只能用羊膜穿刺术）。经皮、脐带取血抽取的纯胎儿血可供检测胎儿血液系统疾病，先天性免疫缺陷，单基因遗传病等。胎儿镜检查可用来诊断大疱性表皮松懈症及某些皮肤疾病，还可以进行肝活检。

第二节　遗传病的治疗

遗传病由于发病机制不同，治疗方法也因此不同。染色体病不仅没有办法根治，改善症状也很难，个别性染色体异常，如 Klinefelter 综合征早期使用睾酮，真两性畸形进行外科手术等，有助于症状改善。多基因病发病中由于环境因素起重要作用，因而药物、外科手术治疗有一定的疗效。目前随着人们对遗传病发病机制的认识逐渐深入，及分子生物技术在医学中的广泛应用，使遗传病的治疗已从常规治疗跨入了基因治疗，为根治遗传病带来了希望。

一、常规治疗原则

（一）手术治疗

如果遗传病已发展到各种临床症状都出现尤其是器官组织已出现了损伤，可对某些遗传病患者进行手术矫正畸形、器官和组织移植来进行治疗。例如，先天性心脏病的手术矫正，肝移植治疗 α_1 抗胰蛋白酶缺乏等。

（二）药物及饮食疗法

遗传病发展到各种症状已经出现时，机体器官已造成一定损害，此时药物治疗主要是对症治疗。或者是针对因代谢过程紊乱而造成的底物或前体物质堆积的情况，进行特殊的饮食疗法或配以药物治疗，以控制底物或前体物质的摄入量，降低代谢产生物的堆积。药物及饮食疗法的原则为禁其所忌、去其所余、补其所缺。

1. 禁其所忌　由于酶缺乏不能对底物进行正常代谢的患者，可限制底物的摄入量以达到治疗的目的。1953 年 Bickle 等首次用低苯丙氨酸饮食疗法治疗苯丙酮尿症患者，收到显效。目前已针对不同的代谢病设计出 100 多种奶粉和食谱。患儿年龄越小，治疗效果越好。

2. 去其所余　由于酶促反应障碍，体内贮存过多"毒物"，可使用多种理化方法将过多的毒物排出或抑制其生成。可用促排泄剂、螯合剂、代谢抑制剂、平衡清除法、换血或血浆过滤等方法减少体内多余的毒物，以减缓症状。

3. 补其所缺　根据某些遗传病的病因，给患者针对性地补充某些成分。

如对某些因 X 染色体畸变引起疾病的女性患者，可补充激素；对分子病及酶病的患者补充相应蛋白质等。

二、基因治疗

基因治疗（gene therapy）是指将正常基因植入靶细胞代替遗传缺陷的基因，或关闭、抑制异常表达的基因，以达到预防和治疗疾病目的的一种临床治疗技术。基因疗法是 20 世纪的一项重大发现，

至今全世界已有2000多名患者作过基因治疗试验，虽然还缺乏令人信服的治疗效果，但基因疗法作为医学界的一项崭新的、划时代的变革，已经引起全世界研究者的重视。

（一）基因治疗的原理与策略

1. 原理　脱氧核糖核酸（DNA）是遗传的物质基础。基因是DNA分子的一个片段，表达产生特异蛋白质，发挥正常的生理功能，从而来维持正常的生命现象。遗传病的根源即在于基因异常，那么若对异常基因给予纠正，就可以使疾病获得根治。基因疗法正是基于这种思考而产生的。

2. 策略　根据宿主病变的不同，基因治疗的策略也是多种多样的，概括起来大致有以下几种：

（1）基因修复：原位修复有缺陷的基因，使其在质和量上均能得到正常表达。目前在技术上似乎还无法做到。

（2）基因代替：指去除整个变异基因，用有功能的正常基因取代之，使致病基因得到永久的更正，目前尚无法做到。

（3）基因抑制和（或）基因失活：导入外源基因除去干扰，抑制有害的基因表达。

（4）基因增强：是指将目的基因导入病变细胞或其他细胞，目的基因的表达产物可以补偿缺陷细胞的功能或使原有的功能得到加强。这一方案最适宜隐性单基因疾病的治疗，目前所作的基因治疗均属此类。

（5）重新开放已关闭的基因：目的在于促使有类似功能的基因表达，以超过或代替异常基因的表达。例如通过去甲基化使已关闭的γ珠蛋白基因重新开放，合成HbF（$\alpha_2\gamma_2$），用以治疗β地中海贫血症。

基因治疗遗传病的策略有多种，但总的来说不外乎是基因的修饰和操作。而就基因转移的受体细胞不同，基因治疗又有两种途径，即生殖（种系）细胞基因治疗和体细胞基因治疗。生殖细胞基因治疗是将正常基因转移到患者的生殖细胞（精细胞、卵细胞和早期胚胎），使其发育为正常个体，显然这是最理想的方法，实质上，这种靶细胞的遗传修饰尚无实质性进展。体细胞基因治疗，指以体细胞为受体细胞，将目的基因转移到体细胞，使之发挥作用以达到治疗的目的。这种方法的理想措施是将目的基因导入靶细胞内染色体上特定基因座位，进行遗传病的治疗。但目前，对特定座位基因转移还有很大困难。

（二）基因治疗的方法与临床应用

1. 基因治疗的方法

（1）基因转移的方法：基因转移是基因治疗的关键和基础。基因转移的途径有两类：一类是in vivo，称为直接活体转移；另一类为ex vivo，称为在体转移。对于遗传病而言，理想的基因治疗是将遗传物质高效率转移到个体细胞中，并且能整合到细胞基因组中，在细胞中长期表达。但是目前的基因转移方法很难满足理想基因转移方法的全部要求，因此探索理想的基因转移方法是基因治疗的一项重要内容。基因转移方法可分为物理、化学和生物学等方法。

1）物理法

A. 显微注射法：是在显微镜直观下，向细胞核内直接注射外源基因。这种方法一次只能注射一个细胞，工作耗力费时。

B. 电穿孔法：是利用脉冲电场提高细胞膜的通透性，使得细胞膜上形成纳米大小的微孔，可将外源DNA转移到细胞中，但有时也会使细胞受到严重损伤。

C. DNA颗粒轰击：这种方法近年来发展较快，它是利用微小的金、钨等贵金属颗粒将DNA吸附，在高压作用下将DNA伴随金属颗粒高速进入细胞，这种方法能有效地使DNA在活体组织、贴壁细胞和悬浮细胞中表达。

D. 脂质体法：该方法是应用人工脂质体包围外源基因，再与细胞融合，或直接注入病灶组织，使之表达。

2）化学法：将正常基因DNA与带电荷物质和磷酸钙、DEAE-葡聚糖、聚-L-鸟氨酸和聚卤化季铵

盐或与若干脂类混合，形成沉淀的 DNA 微细颗粒，以加强细胞摄取外源 DNA。但此法转移效率比较低，成功率在 $1/10^3 \sim 1/10^2$。

3）生物学法：主要指病毒介导的基因转移，包括 RNA 病毒、DNA 病毒两大类。病毒为载体是当今最有效的转移目的基因的方法，能够用作载体的病毒有 SV40、牛乳头状瘤病毒、单纯疱疹病毒 I 型、巨细胞病毒、腺病毒和反转录病毒等。其中较常用的是反转录病毒和腺病毒。

A. 反转录病毒载体：反转录病毒虽然是 RNA 病毒，但有反转录酶，可使 RNA 转录为 DNA。再整合到宿主细胞基因组。反转录病毒具有高效感染细胞、稳定表达外源基因、宿主范围广泛等优点；但它也有基因容量小、可致细胞恶变等缺点，然而这些缺点可通过人工改建和体外包装而克服。在反转录病毒载体中，最常用的是莫洛尼鼠白血病病毒（murine leukemia virus；Mo-MLO）。

B. 腺病毒（adenovirus，AD）载体：腺病毒属于 DNA 病毒，是一类无胞膜病毒，基因组长约 36kb。AD 载体大多是以 2 型和 5 型 AD 为基础构建而来，其中多缺失 E1 和 E3 区，代之以外源目的基因。AD 载体因为缺失 E1 和 E3 区，自身不能表达包装蛋白，故需要在 293 细胞或 Hela 细胞（含有 E1 和 E3 区）的帮助下才能形成病毒颗粒，从而感染细胞。它具有插入 DNA 较长，不需要正在分裂的靶细胞，可以原位感染和病毒滴度高等优点。另外，腺病毒载体一般不会整合到宿主的基因组中，从而大大减少了插入突变的潜在危险。

应用腺病毒载体已用来作肺、肝、中枢神经系统、内皮细胞、肌细胞和唾液腺上皮细胞的基因治疗实验，对呼吸系统的遗传性疾病和肿瘤，可采用多次滴注法，特别有效方便。目前，这一方法在临床上用于囊性纤维化患者的基因治疗。

另外还有通过受体介导的基因转移、同源重组法进行的基因转移。有资料表明，前者外源基因在活体内的表达只是暂时性的，因此，此种方法的应用前途可能有限。后者如能改进技术，提高重组率，仍是有广阔前景的。

（2）反义核酸疗法：包括反义 RNA 技术和反基因技术两种。前者是将人工合成的反义 RNA 导入靶细胞，与特定 mRNA 分子互补结合，抑制特定基因的表达。近来又有人引入核酶，使 mRNA 特异的分子降解，以抑制特异基因的表达。反基因技术是将一段 DNA 分子导入靶细胞，与 DNA 双螺旋分子的专一序列形成三螺旋 DNA 以阻止基因的转录。

此外，还可将“自杀基因”（如 T_K 基因），多药抗药性基因、抑癌基因（如 *P53*）等基因导入靶细胞进行疾病的治疗。

2. 临床应用　目前发现的遗传病已经达到 8000 多种，然而由于受种种因素的限制，迄今为止，只有 20 种遗传病被列为基因治疗的主要对象，其中部分疾病研究已进入了临床试验阶段。（表 12-2）

表 12-2　目前临床试用的基因治疗遗传性疾病

疾　病	传递的基因或产物	靶细胞或组织	载　体
α-抗胰蛋白酶缺乏症	α-抗胰蛋白酶	呼吸道	脂质体
慢性肉芽肿	$p47^{PHOX}$	骨髓细胞	反转录病毒
囊性纤维化	囊性纤维化跨膜调节蛋白	呼吸道	腺病毒、脂质体腺伴随病毒
家庭性高胆固醇血症	低密度脂蛋白受体	肝细胞	反转录病毒
范可尼综合征	互补组 C 基因	造血祖细胞	反转录病毒
高雪病	葡萄脑苷酯酶	周围血细胞或造血干细胞	反转录病毒
亨特综合征	艾杜糖醛酸-2-硫酸	淋巴细胞	反转录病毒
腺苷脱氨酶缺乏引起的免疫缺陷病	腺苷脱氨酶	淋巴细胞	反转录病毒

(1) ADA缺乏症:腺苷脱氨酶(ADA)的缺乏可使T淋巴细胞因代谢产物的累积而死亡,从而导致严重的联合性免疫缺陷症(SCID)。大约25%的SCID患者是由于ADA缺乏引起的。近年来,ADA缺乏症已成为遗传病基因治疗的首选疾病。主要集中在用反转录病毒载体将基因*ADA*转移到离体细胞中,均能较好地表达外源基因*ADA*,而且可以在小鼠体内长期表达,达到治疗水平。1990年9月美国Blease小组对一4岁ADA缺乏症(AR)儿童进行了世界首次基因治疗临床试验:向患者体内输注遗传修饰的自身T细胞,使患者的症状得到改善。1999年,法国学者Cavazzana-Calvo等在造血干细胞的水平上,对两名患ADA缺乏症(XR)的婴儿(分别为11个月和8个月龄)进行基因治疗,经10个月的随访,情况良好,T、B细胞和天然杀伤细胞的数目以及免疫系统均正常。这可以称得上是人类历史上第一次真正意义上的基因治疗。

(2) 血友病B:血友病B是X连锁隐性遗传病,患者凝血Ⅸ因子缺乏(相应基因定位在Xq27.1-q27.2),临床表现为出血、凝血时间异常。我国复旦大学薛京伦教授等尝试将Ⅸ因子基因插入反转录病毒载体并转移到患者皮肤成纤维细胞中,获得表达;随后把这些转染的细胞再送回到患者体内以期达到治疗目的。结果显示,患者体内表达Ⅸ因子,其临床症状也有所改善,但其长期疗效还有待观察。

(3) 囊性纤维化(cystic fibrosis,CF):以病毒为载体研究进行的囊性纤维化基因治疗在体外培养细胞中获得了成功,即转入的正常基因纠正了CF的缺陷。但在整体水平上CF患者的主要病理改变发生于呼吸系统的上皮细胞,而如何将基因导入这些细胞并回到其功能位置,则是一件十分棘手的事。

(三) 基因治疗存在的问题

基因治疗应用于遗传病的治疗只有较短的时间,虽然它对基础医学或临床医学的发展均起到了推动作用,但应清醒地看到,目前的基因治疗尚存在许多问题。

1. 稳定性　目前在动物实验和人的临床治疗中,基因通过媒介进入靶细胞后,基因表达不稳定,甚至不表达。靶细胞在复制时,新基因可能被丢失。这可能因为:①基因转录系统不稳定;②形成不正确的信息RNA;③基因表达的控制因素复杂;④靶细胞寿命短,产生毒素等。现已有许多实验室正在研究寿命较长的靶细胞,如造血干细胞和骨髓前体细胞,法国已有成功的例子。

2. 安全性问题　应用病毒载体进行基因治疗的安全性,已引起广泛重视,主要预防下列问题:一是感染,二是有益基因的丢失;三是诱发癌变。

3. 导入基因的高效表达问题　迄今所有导入细胞的目的基因表达率都不高,目前不少实验室正在研究将高效的启动子构建入反转录病毒载体。

4. 免疫性　临床治疗有时需要多次操作,使机体产生免疫反应,排斥携带基因的病毒或靶细胞,给进一步治疗造成困难。为了减少免疫反应,有些研究者将一种基因组合到若干不同的腺病毒中,这样有可能避免免疫反应发生。另外,尽可能多地将与免疫有关的病毒基因删除也是解决问题的方法。

5. 伦理问题　对遗传病进行基因治疗要取得社会的理解与配合。为此,要宣传基因治疗的科学性与安全性以及对人类健康的重要性,以提高人们的认识,同时要建立并完善医疗法制与措施。

基因治疗是遗传病治疗的一种崭新手段,它正逐渐被人们所接受。基因治疗所蕴藏的巨大潜力有力地证明,遗传病一定是可以治疗的。科学家们建议,今后要加强基础研究,解决目前基因治疗遗传病的一些问题,再进行大规模的临床试验。我们深信,只要坚持严格的科学态度,设计精细的临床研究,基因治疗在21世纪必将获得令人振奋的成果。

第三节　遗传病的预防

近年来,我国遗传性疾病比例上升,预防遗传病性疾病发病的问题显得日益重要。遗传性疾病目前多无有效的治疗方法,因此开展好遗传病的预防十分重要。国际上所采用的遗传咨询、产前诊断及

遗传筛查三结合的方法，有效地降低了各国常见遗传病的发病。我国遗传病的预防也从这三方面进行。为了更好地了解各地区的遗传病种以针对它进行有效的预防，控制其在一些家庭中的发生和群体中的流行，遗传病的登记和随访、遗传保健也是遗传病预防不可缺少的方法。本节就遗传病的筛查、咨询、登记和随访、保健等方面对遗传病的预防进行论述。

一、遗传筛查

遗传筛查（genetic screening）是研究群体各成员某一位点基因类型的一项普查。通过筛查，可及早发现有致病基因的个体；同时，获得的数据有利于遗传病的发病规律和流行特点的研究。遗传筛查包括出生前筛查、新生儿筛查、携带者检测三个方面。

（一）出生前筛查

出生前筛查是诊断胎儿有无遗传性疾病的过程。出生前筛查是生化遗传、细胞遗传、分子遗传和临床实践相结合的产物，具有很强的实际应用价值。近年来发展很快，国际上目前除染色体病外，还有约 100 多种遗传病可作出生前筛查。在遗传咨询的基础上，对有高风险的妊娠经产前诊断判断胎儿患病时，可终止妊娠，防止患儿的出生，这是预防严重遗传病患儿出生的有效手段。

可进行出生前筛查的遗传病包括以下几类：染色体病；特定酶缺陷所致的遗传性代谢病；可进行 DNA 检测的遗传病；多基因遗传的神经管缺陷（neural tube defect，NTD）；有明显形态改变的先天畸形。

（二）新生儿筛查

目前，有一些遗传病已有有效疗法，若能在新生儿阶段明确该种疾病的诊断，在患儿出现不可逆的损伤前得到治疗则可防止临床症状的出现。

新生儿筛查（neonatal screening）选择的病种应考虑下列条件：发病率较高；有致死、致残、致愚的严重后果；有较准确而实用的筛查方法；筛出的疾病有办法防治；符合经济效益。

有些国家已将新生儿筛查列入优生的常规检查，筛查的病种达 12 种。我国这项工作刚起步，列入筛查的疾病有苯丙酮尿症（phenyketonuria，PKU）、先天性甲状腺功能低下、G6PD 缺乏症（南方）。

苯丙酮尿症是由于苯丙氨酸羟化酶缺乏所致的一种代谢病，它是一种常染色体遗传病。此症临床表现严重，但如果能在新生儿期发现，就可以通过饮食控制等措施防止或减缓症状的出现和发展。Guthrie 细菌抑制法是国际上普遍采用的筛查方法。此方法是将干血片置于含有与苯丙氨酸结构相似的细菌抑制剂 β-2-噻吩丙氨酸（β-2-thienylalanine）的培养板上，孵育过夜后，观察干血片周围的细菌生长环。在正常标本，枯草杆菌因受培养板中抑制剂的作用而不能生长或出现较小的生长环。若血液中苯丙氨酸浓度增高时，则细菌生长就不受抑制，培养板上出现大菌环。筛查所得阳性个体均应采静脉血作苯丙氨酸及酪氨酸测定。一旦明确诊断，应立即进行饮食控制，减少苯丙氨酸的摄入。

在新生儿筛查工作中应强调的是：必须有完善的遗传病登记；有敏感、准确的筛查方法；筛查出的新生儿应送到遗传咨询中心，经有经验的临床生化遗传专家进一步确定诊断；对确诊的患儿提出治疗方案并定期随访。

（三）携带者筛查

遗传携带者（genetic carrier）指表型正常，但带有致病遗传物质（致病基因或染色体畸变），能传递给后代使之患病的个体。一般包括：带有隐性致病基因的个体（杂合子）；带有平衡易位染色体的个体；带有显性致病基因而暂时表达正常的顿挫型或迟发外显者。

携带者筛查（carrier screening）是指当某种遗传病在某一群体中有高发病率，为了预防该病在群体中的发生，采用经济实用、准确可靠的方法在群体中进行筛查，筛出携带者后则进行婚育指导，即可达到预期目标。携带者筛查对遗传病的预防具有积极意义，表现在：人群中许多隐性遗传病的发病率较低，但杂合子的比例却相当高，如遇到两个携带者婚配，及时检出这些隐性基因携带者，进行婚育指导，意义很大；染色体平衡易位者可有较大比例出生死胎或染色体异常患儿，如母亲是染色体 14/21 的平衡易位携带者，其子女中，正常儿、携带者和患儿各占 1/3，一部分缺少一条染色体的胎儿不能存

活而中途流产，所以及时检出有助于对该病的确诊和发病风险的推算，也便于进行遗传咨询和指导；对显性遗传病的携带者，如能及时检出，更可以预先控制发病的诱因或中间环节，防止发病或阻止病情进展，意义更大。

Tay-Sachs 病是一种缺乏 N-乙酰-β-D-氨基半乳糖苷酶（已糖苷酶 A）α 亚单位的常染色体隐性遗传病。该酶的缺乏可引起严重的神经退行性疾病，患者在 2～4 岁死亡，无法治疗。北欧犹太人人群中 3% 为 Tay-Sachs 病携带者，因此筛查在这个人群中进行。测量血白细胞中己糖苷酶 A（hexA）的活性来确定携带者的情况。产前诊断对 Tay-Sachs 病是有用的，以便后代中有风险的夫妇对怀有的受累胎儿终止妊娠。在过去的 25 年中该筛查已使北欧犹太人该病的发病率大大降低。

二、遗传咨询

遗传咨询（genetic counseling）是由临床医师和遗传学工作者解答遗传病患者及其亲属提出的有关遗传性疾病的病因、遗传方式、诊断、治疗及预防等问题，估计患者的子女再患某病的概率，并提出建议及指导，以供患者及其亲属参考。遗传咨询的意义在于：减轻患者身体和精神上的痛苦，减轻患者及其亲属的心理压力，帮助他们正确对待遗传病、了解发病概率，采取正确的预防、治疗措施；降低人群遗传病的发生率，降低有害基因的频率，及减少传递机会。

（一）遗传咨询的对象和步骤

1. 遗传咨询的对象　需要进行遗传咨询的人群有两种类型，一类是主动前来咨询的，这类人一般是因为一家中几个成员都患有某种遗传病，可以意识到自己或子女有患病风险；另一类是由咨询医师根据“遗传登记”经其宣传、解释后，被动前来咨询的，此类人群一般属于散发的遗传病家庭，由于看不到发病的家族聚集，未注意到患病或再发风险。归纳起来有以下情况者应进行遗传咨询：①夫妇一方患有某种遗传病，需要给予生育指导；②一对夫妇生了一个遗传病患儿，询问再发风险者；③一对夫妇婚后多年不育或妻子出现不明原因的习惯性流产，要求从遗传角度寻找不育或流产的原因；④婚前或婚后了解到家属中有遗传病患者，担心子代是否也会患此遗传病；⑤近亲婚配的夫妇，要求给予生育指导；⑥家庭成员中得了病因不明的疑难杂症，要求肯定或排除遗传病的可能性。

2. 遗传咨询的步骤　遗传咨询过程中，咨询医师是起主导作用的，对咨询者来说，是一个解疑求助的简短的教育过程，然而，遗传咨询必须贯彻非指令性遗传咨询的原则。另外咨询医师要根据患者或患儿父母心理学上的变化进行必要的开导，使他们理智地面对现实，才能使咨询达到良好的效果。遗传咨询可遵循下列程序：

（1）认真填写病例：填写详细的遗传咨询病历，并妥为保存，备后续咨询用。

（2）对患者作必要的体检，作出诊断：根据患者的症状和体征，建议患者作辅助性检查及必要的、有针对性的实验室检查。有时这类检查还需扩展到其一级亲属。一般在第二次或第三次咨询作出初步诊断。在判定是否为遗传病时，咨询医师要排除一些干扰性因素，以明确诊断。

（3）对再发风险的估计：由于部分遗传病是致残、致愚的，甚至是致死的，故应对那些需求生育第二胎的咨询者作出再发风险的估计（估计方法详见下面内容）。

（4）与咨询者商讨对策：包括劝阻结婚、避孕、绝育、人工流产、人工授精、产前诊断、积极改善症状等措施。此时应由咨询者选择由咨询医师提出的方案。

（5）随访和扩大咨询：为了明确咨询者提供信息的可靠性，观察遗传咨询的效果和总结经验教训，有时需要对咨询者进行随访，以便改进工作。如果从全社会或本地区降低遗传病发病率的目标出发，咨询医师还应主动追溯家属中其他成员是否患有该病，特别是查明家属中的携带者，这样可以扩大预防效果。

（二）遗传病再显危险率的估计

再显危险率的估计是遗传咨询的核心内容，也是遗传咨询门诊有别于一般医疗门诊的主要特点。

再显危险率又称为复发风险率，是指曾生育过一个或几个遗传病患儿，再生育该病患儿的概率。现在这一概念已经扩大到凡有信息可导致一对夫妇再生育患儿（包括第一胎）的概率，但这一情况称患病风险较适当。

再发风险的估计一般遵循下列原则：染色体病和多基因病以其群众发病率为经验危险率，只有少数例外。单基因病则根据孟德尔规律作出再发风险的估计，介绍如下：

1. 染色体病再发风险的估计　染色体是遗传物质的载体，其数目和结构的相对稳定是个体基因组的完整、结构和功能表达正常的保证，更是维持生物遗传性状相对稳定的基础。染色体病一般均为散发性，其畸变主要发生在亲代生殖细胞的形成过程中，因此再发风险率实际上就是经验危险率或称群体发病率。临床上很少见到一个家庭中同时出现 2 个或 2 个以上染色体病患者。

然而，也有一些例外的情况，如双亲之一为平衡易位携带者或嵌合体，子代就有较高的再发风险率。下面以易位型 Down 综合征为例说明之。假如父亲或母亲的染色体核型是 45，XX(XY)，-14，-21，+t(14q21q)，由这种核型所产生的生殖细胞与正常生殖细胞形成受精卵时，可产生 6 种不同的核型（图 12-5）。其中 21 单体型和 14 单体型是致死的；14/21 易位型 14 三体综合征也很少能成活；剩下的要么是 14/21 易位型 21 三体综合征，要么是平衡易位携带者，要么是正常个体，且理论上各占 1/3。但实际上 14/21 易位型 21 三体综合征的出生率要低于上述理论值，原因可能与流产有关。

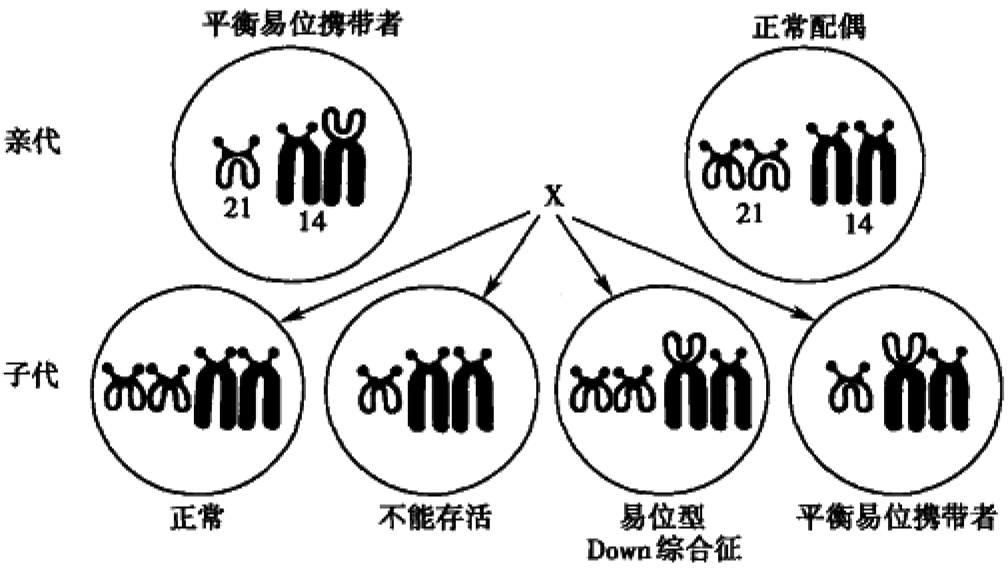


图 12-5　14/21 易位型 Down 综合征家系子代的各种可能性

另外，母亲是平衡易位携带者，其子代风险要高于父亲是平衡易位携带者，原因可能在于母亲每月只排出 1 个卵细胞，不像精子存在机遇。表 12-3 示 21 三体综合征的再发风险率。

表 12-3　21 三体综合征的再发风险率

核型			再发风险率
患者	父亲	母亲	
D/G 易位型	正常	携带者	0.10～0.15
	携带者	正常	0.05
21/21 易位型	正常	携带者	1.00
	携带者	正常	1.00
21/22 易位型	正常	携带者	0.10～0.15
	携带者	正常	0.05
21 三体型	正常	正常	0.01
易位型或嵌合型	正常	正常	小

还有应注意的是大多数三体综合征的发生与母龄呈正相关，即随着母亲年龄增大，三体综合征的再发风险率也随之增大。有人推测，这可能是由于35岁以上的妇女的卵巢开始退化，从而导致卵细胞形成过程中高发染色体不分离之故。

2. 单基因病再发风险率的估计　单基因病再发风险率可根据家系咨询提供的信息，并按孟德尔遗传规律加以估计。如果所获信息能肯定亲代的基因型，那么子代的再发风险率可按单基因不同遗传方式的传递规律加以估计。如果所获信息还不足以肯定亲代的基因型，那么子代的再发风险率可按 Bayes 逆概率定理加以估计。

（1）亲代基因型已推定时再发风险率的估计：一对夫妇的基因型如能通过他们本身或他们的父母、子女所患的遗传病加以推定，那么根据此遗传病的遗传方式就能计算出子代的再发风险率。

1）常染色体显性遗传病：此类疾病的显性纯合子一般均在胎儿期死亡或幼年死亡，极少数能活到成年并有生育能力。因此，能结婚并生儿育女的主要是杂合子患者。夫妇一方患病时，子代每胎再发风险率是1/2；夫妇方均为患者时，子代再发风险率为3/4；夫妇双方均正常时，子代再发风险率是0。

2）常染色体隐性遗传病：此类疾病的患者均为隐性纯合子。因此，一对表型正常的夫妇生了一个病孩，此时即可推定这对夫妇双方均为杂合子，他们子代再发风险率是1/4，表型正常的子代是杂合子的可能性为2/3，即(2/4)/(1/4＋2/4)＝2/3，完全正常的机会是1/4；如夫妇一方为患者，另一方为显性纯合子，此时子代不会发病，但全部是杂合子；如夫妇一方为患者，另一方为杂合子时，子代发病机会是1/2，携带者的机会也是1/2。需要注意的是遗传异质性现象，如白化症夫妇或先天性耳聋夫妇生育了正常子代，这是因为这对夫妇的致病基因不在同一位点上，造成子代为双重杂合子(double heterozygote)，但不构成隐性纯合子。图12-6示常染色体隐性遗传病家系中表型正常亲属是杂合子的概率。人体基因在同源染色体上成对存在，子代得到一个亲代一对等位基因中某个特定等位基因的机会是1/2。如 AA×Aa 婚配，子1代也是 Aa 杂合子的机会是1/2，子2代也是 Aa 杂合子的机会是(1/2)(1/2)＝1/4。同理，Aa×Aa 婚配，前已述及表型正常子1代的杂合子机会为2/3，子2代也是 Aa 杂合子的机会就应为(2/3)(1/2)＝1/3。

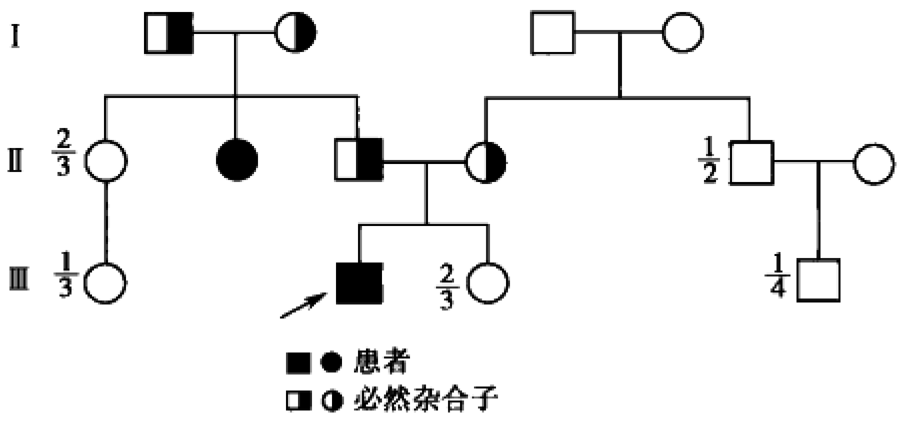


图12-6　常染色体隐性遗传病家系表现型正常亲属的杂合子概率

3）X连锁显性遗传病：此类疾病的发病率男女有别。当丈夫患病、妻子正常时，他们的儿子全部正常，而女儿全部是杂合子患者；当妻子有病、丈夫正常时，他们的儿子和女儿的发病机会均为1/2；当夫妇双方均为患者时，女儿全部得病，而儿子仅有1/2机会得病。

4）X连锁隐性遗传病：此类疾病，女性患者为隐性纯合子，男性患者为半合子。在丈夫患病、妻子正常时，儿子全部正常，女儿全部是杂合子；在妻子是患者、丈夫正常时，儿子全部患病，即再发风险率为1，女儿全部是杂合子；在妻子为杂合子、丈夫正常时，儿子得病机会是1/2，女儿得病机会为0，但女儿有1/2机会成为杂合子；在丈夫为患者、妻子是杂合子时，儿子得病的机会是1/2，女儿得病机会也是1/2。图12-7示X连锁隐性遗传病家系男性亲属再发风险率和女性亲属杂合子概率。

（2）亲代基因型未能推定时再发风险率的估计：如果双方或一方的基因型未知，这时则要利用家系资料或其他有关数据，用 Bayes 逆概率定理来推算。Bayes 定律是概率论的基本定律之一，是一种确认两种相互排斥事件相对概率的理论。按照 Bayes 理论，遗传咨询中的概率计算包括下述几个层次：

1）根据遗传规律算出携带者的概率，称为前概率(prior probability)。对同一遗传病的每一家系，每一组合的前概率都是固定不变的。

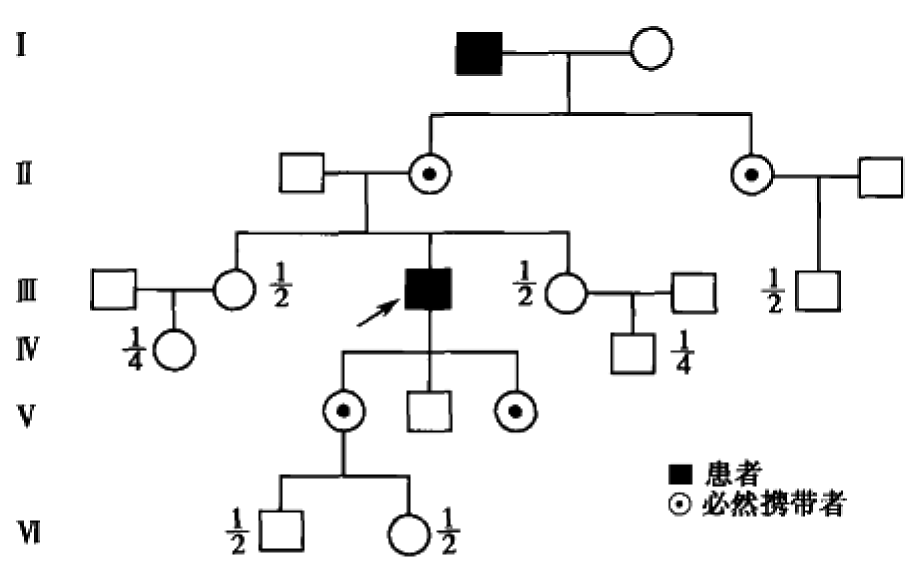


图 12-7　X 连锁隐性遗传病家系男性亲属再发风险率和女性亲属杂合子概率

2）从咨询者的子代发病情况等条件，算出条件概率（conditional probability）。

3）将前概率和条件概率相乘，算出各自的联合概率（joint probability）。

4）将所有联合概率相加作为分母，将每项联合概率作为分子，即可得出后概率（posterior probability），又可称为总概率。由于后概率除前概率外，还包括该家系的其他信息，所以数据更为准确。

A. X 连锁隐性遗传：按孟德尔分离律，男性患者（半合子），全部女儿为杂合子；若女性为杂合子，则子女中携带此基因的概率各为 1/2；但男性若获得此基因（半合子）则可发病，女性为携带者。某些常见的 X 连锁遗传病，女性杂合子不易测出，有时出现男方患病，儿子也罹患的情况（母亲为未测出的杂合子）。如果女性杂合子完全不能检测或测出机会很少时，可用 Bayes 定律来推测，例如图 12-8 是一例假肥大型肌营养不良（DMD）的系谱。从$Ⅱ_1$ 为患者、$Ⅱ_2$ 为肯定携带者（已生育一个患儿$Ⅲ_1$）来分析，$Ⅰ_2$ 是杂合子无疑。问题是$Ⅱ_3$ 是否为杂合子携带者？由于$Ⅰ_2$ 是杂合子，故$Ⅱ_3$ 是携带者的前概率是 1/2，她不是携带者的前概率也是 1/2。她已生了 3 个儿子都正常，如果她是携带者，而 3 个儿子正常，则条件概率是$(1/2)^3=1/8$；如果她不是携带者，3 个儿子正常，则条件概率为 $1^3=1$。这样$Ⅱ_3$ 是携带者的合并概率为 $1/2\times1=1/2$。由此她是携带者的后概率为：$(1/16)/(1/16+1/2)=1/9$，即由于$Ⅱ_3$ 已生育了 3 个正常男孩，她是携带者的概率由 1/2 降至 1/9，而不是携带者的可能性是 8/9。

B. 常染色体隐性遗传：按一般孟德尔定律计算再发风险，即 1/4。如生育一个以上患儿再发风险仍为 1/4。但有时要作具体分析。例如：

a. 一对表兄妹拟结婚，而他们的亲属中有一种常染色体隐性遗传病患者，如系谱图 12-9 所示：$Ⅲ_3$ 为患者，$Ⅲ_2$ 外表正常，故有 2/3 的机会为杂合子，考虑到 $Ⅰ_1$ 和 $Ⅰ_2$ 一方为杂合子的可能性较大，在这种情况下，$Ⅱ_1$ 和 $Ⅲ_1$ 为携带者的概率分别为 1/2 和 1/4。$Ⅲ_1$ 和 $Ⅲ_2$ 结婚生育第一胎为患儿的概率为 $1/4\times2/3\times1/4=1/24$。若表兄妹结婚亲属中未发现有某种常染色体隐性遗传病患者，则所生子女患病风险将根据该病在人群中发病率和近婚系数推算。

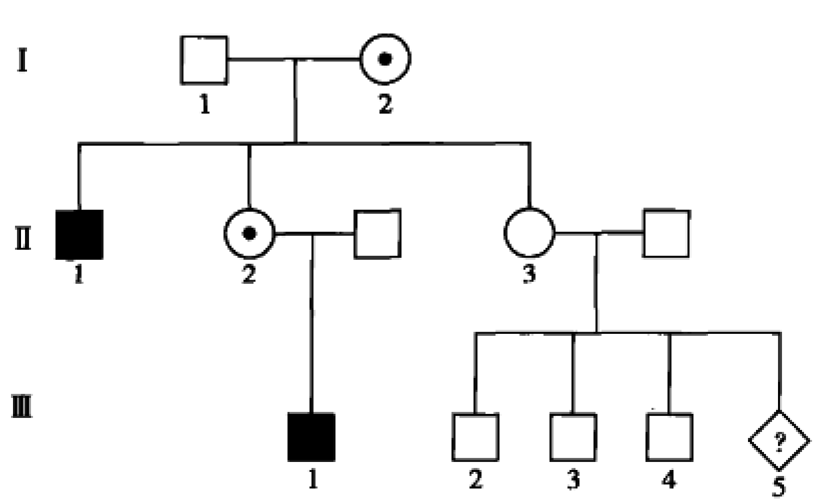


图 12-8　一例假肥大型肌营养不良系谱

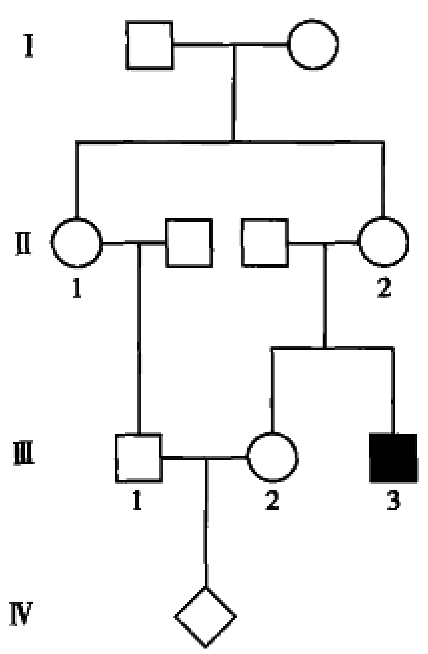


图 12-9　一例常染色体隐性遗传病系谱

b. 双方为某种常染色体隐性遗传病患者,结婚后生子女患该病的风险如何？按孟德尔分离律,他们的子女将全部为患者。但事实上并非如此。据一项观察,6对聋哑夫妇结婚,他们中仅一对为全部子女聋哑。这是由于疾病的遗传异质性造成,表型相同的患者,由于存在不同的基因型,所以他们的后代可以是不同基因的双重杂合子,不表现症状。

C. 常染色体显性遗传:在一般情况下,常染色体显性遗传病的再发风险为1/2,即使已生育一个或多个患儿,再发风险仍为1/2,但在下列情况下这一概率会有所变动。

a. 迟发显性:显性症状要在若干年后才表现。例如Huntington舞蹈病(Huntingtons chorea,HC)是一种常染色体显性遗传病,携带此基因的杂合子一般发病较迟。图12-10中咨询者$Ⅱ_2$的父亲为患者,母亲正常,则咨询者获得致病基因的概率为1/2,未获得此基因的概率亦为1/2。随着咨询者年龄的增长而不出现症状,表明她实际上获得致病基因的概率日益减少,而未获得致病基因的概率则越来越大。据统计,本病于30岁前发病约占1/3。咨询者为30岁时,如果她携带有致病基因,她尚有2/3的机会发病,而如果她未携带此基因将不会发病。此咨询者患病的总概率或后概率可计算,如表12-4。

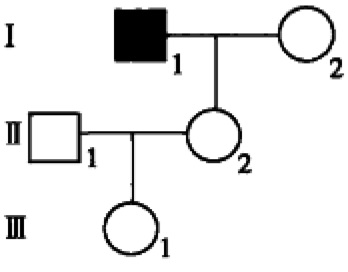


图12-10 Huntington舞蹈病家系

表12-4 Huntington舞蹈病患者儿子30岁时为携带者的概率

	获得致病基因的概率	未获得致病基因的概率
前概率	1/2	1/2
条件概率(30岁时)	1 - 0.33 = 0.67	1
合并概率	1/2 × 0.67 = 0.33	1/2 × 1 = 0.5
后概率	0.33/(0.33 + 0.50) = 0.4	0.50/(0.33 + 0.50) = 0.6
携带者风险	0.40	

从表12-4可以看出:由于咨询者30岁仍未表现症状,故她是患者的风险由50%降至40%,如果她40岁仍未发病,她患病的风险还会降低。

b. 外显不完全:在显性遗传病中,如有外显不全的情况(即外显率低于100%),此时,子女患病概率为1/K(K为外显率)。例如视网膜母细胞瘤的遗传方式为常染色体显性遗传,其外显率为70%,按公式计算,生育患儿的概率为1/2×0.70=0.35(35%);携带者(这里指携带显性基因而未表现病症的个体)的概率为1/2(1-K),即0.15(15%)。又例如父亲为患者,女儿外表正常,假设此基因的外显率为90%,试问女儿的下一代患病风险有多大？此时有几种可能:女儿未携带此基因,前概率为1/2;女儿携带此基因但未表现,则其合并概率为:1/2(前概率)×1/10(条件概率)=1/20;她作为一个携带者的总概率(后概率)为:(1/20)/(1/20+1/2)=1/11,即此女儿为携带者的概率只有1/11,而她的第一个子女的患病风险为:1/11×1/2×9/10=9/220,即约4%。

3. 多基因病再显危险率估计　多基因遗传病是遗传因素和环境因素共同作用所致,故不能像单基因遗传病那样通过分离律和自由组合率来确切地算出其再发风险率,而只能通过群体发病率和家系中受累者的多少来加以估计,这种估计概率称为经验危险率(empirical risk)。表12-5列举了一些常见多基因遗传病的经验危险率。另外,根据Edwards公式,如某种多基因遗传病的群体发病率在0.1%~1%,遗传度在70%~80%,这时患者一级亲属的发病率(q_r)为群体发病率(q_g)的平方根,即$g_r=\sqrt{q_g}$从此公式可知,群体发病率越低,一级亲属中再发风险率相对地越高。

多基因遗传病的发病还具有下列特点,即亲缘关系越近,再发风险率越大;家系中患者数越多,再发风险率也越大;再有该病的遗传度越高,一级亲属的再发风险率也越高(可参阅第六章多基因遗传表6-4)。近年来,由于一些实用的多基因遗传数学模型的相继建立,加上电子计算机的逐步普及,使多基因遗传病再发风险率的估计更趋准确。同时,由于对多基因遗传基础中主基因(major gene)的研

究逐步深入，可望对此类疾病的研究也进入分子水平。在不久的将来，对多基因遗传病再发风险率的估计必将会有较大的突破。

表 12-5 常见多基因遗传病的经验危险率(%)

疾 病	群体发病率	男:女	父母正常两个孩子受累	一个亲代和一个孩子受累	一个亲代和两个孩子受累
无脑儿	0.20	1:2	2	—	—
腭裂	0.04	2:3	2	7	15
腭裂+唇裂	0.10	3:2	4	4	10
畸形足	0.10	2:1	3	3	10
多发先天性心脏病	0.50	—	1~4	1~4	—
早发型糖尿病	0.20	1:1	3	3	10
髋关节脱位	0.07	1:6	4	4	10
癫痫(特发性)	0.50	1:1	5	5	10
巨结肠(Hirschsprung 病)	0.02				
男性先证者		4:1	2	—	—
女性先证者			8	—	—
躁狂抑郁型精神病	0.40	2:3	5~10	5~10	—
智力障碍(特发性)	0.30~0.50	1:1	3~5	—	—
幽门狭窄	0.30	5:1			
男先证者			2	4	12
女先证者			10	17	38
精神分裂症	1~2	1:1	14	16	—
脊柱侧凸(特发性青年型)	0.20	1:6	7	5	—
脊柱裂	0.30	2:3	4	—	—

三、遗传登记和随访

(一) 遗传登记

遗传医学中心为了控制该地区某遗传病的发生，在遗传咨询的基础上必须进行遗传登记(genetic register)。它是指某一地区遗传保健服务机构对该一地区某些严重遗传病的家系进行登记的一项工作。遗传登记的类型有很多，根据不同的目的可分为以下几种：①临床遗传登记，目的在于观察某些遗传病的发病过程及不同治疗手段的效果等，又可使一些新的诊断和治疗手段及时提供给先证者亲属以便早期发现、诊断和治疗；②遗传流行病学登记，目的是为确定某个群体中遗传病的发病率和流行规律，以便正确估计遗传因素、环境因素在遗传发病中所起作用的大小，从而促进遗传病的预防工作；③跟踪遗传登记，目的是估计遗传咨询或产前诊断的效果，同时也可以对一个地区的遗传保健工作作出评估；④预防性遗传登记，通过对高风险产妇进行遗传咨询和产前咨询，减少遗传病的发病率和遗传病的负荷。遗传登记的适应证一般是：群体发病率较高，症状比较严重，且大多发病较晚，又无很好治疗手段的遗传病。表 12-6 是遗传登记和部分适应证。

遗传登记的内容应包括个人病史、发育史、婚育史、生育次数、亲属病情、系谱绘制、风险个体、近亲婚配、资料的统计整理等内容。遗传登记中，必须注意一个重要问题：遗传登记是为遗传病家系服

务,贮存的数据均为有关家系的隐私,属于保密范围。

表 12-6 遗传登记的部分适应证

类 型	遗 传 病	类 型	遗 传 病
染色体病	易位型 Down 综合征		神经纤维瘤病
常染色体显性遗传病	多囊肾病		色素性视网膜炎
	Huntington 舞蹈病	X 连锁隐性遗传病	Duchenne 肌营养不良症
	Marfan 综合征		Becker 肌营养不良症
	强直性肌营养不良		血友病 A 和 B
	结肠息肉症		色素性视网膜炎

(二) 遗传随访

遗传随访(genetic follow-up)是对已确诊的遗传病患者及其家属作定期的门诊检查或家访,以便动态观察患者及其家属各成员的变化情况,同时给予必要的医疗服务。随访又分短期随访和长期随访两种。

短期随访一般为期 3 个月,目的是保证患者及其亲属理解咨询过程中医师所提供的信息,懂得减少或避免子代再显危险率的方法,以及医师在第一次咨询过程中未表明的一些问题。

长期随访是指保持与患者及其亲属的长期联系,时间可持续 10 年以上,以便及时发现患者及其家属的变动情况,包括地址变更、婚姻状况、生育情况、患者表现型的变化以及新病例的发病情况等。如果家系中有高风险成员婚配,应及时进行婚姻和生育指导。另外,如有新的诊断或治疗措施问世,也可及时提供给家系各成员。长期随访还有一个优点,即可以与患者及其家属长期沟通,以便及时发现和解决他们的一些心理问题,使他们能正确对待疾病,勇敢地面对现实。一般来说,对遗传登记的家系均应进行长期随访。

四、遗传保健

遗传保健(genetic health care)是遗传医学的一个组成部分,它不仅为遗传病患者提供服务,更重要的是为遗传病家系成员和人群中的遗传病高风险对象提供医学遗传服务。它要为遗传病患者提供最好的现代医学处理,为遗传病家系成员和人群中的遗传病高风险者提供遗传咨询,通过婚前咨询、群体筛查、杂合子检出、出生前诊断、症状前诊断等各个环节的措施预防出生遗传病患儿,尽可能保证遗传病家系成员的健康。

遗传保健首先是防止环境污染(废气、废水、废渣放射),其次是纠正不良的生活习惯(吸烟、酗酒、食用不当食物等)。遗传保健是提高人口素质的根本所在,最终对人类社会产生巨大的推动作用。

Summary

Clinical genetics is the important branch of medical genetics involved with the diagnosis, treatment and prevention of genetic disorders affecting individuals and their families. Firstly, diagnosis of genetic disorders is based on taking an accurate history and performing clinical and experimental examination. The history and examination should focus on aspects relevant to the genetics, such as drawing a pedigree, estimation of risk and gene diagnosis. Secondly, genetic disorders may be amenable to treatment, either symptomatic or potentially curative. Treatment may range from conventional drug or dietary management and surgery to the future possibility of gene thera-

py.Finally, genetic counseling is an integral part of the genetic testing process. The extent of the counseling and the issues to be addressed depend upon the type of test being offered, which may be diagnostic, presymptomatic, carrier or prenatal testing. The ultimate aim of genetic counseling is to help families at increased genetic risk to live and reproduce as normally as possible.

（孙树汉）

参考文献

1. 孙树汉. 遗传与疾病. 北京：人民卫生出版社，2009
2. Gelehrter TD, Collins FS, Ginsberg D. 医学遗传学原理. 孙开来译. 北京：科学出版社，2000
3. 孙树汉，胡振林，颜宏利. 染色体、基因与疾病. 北京：科学出版社. 2008
4. 陆国辉，徐湘民. 临床遗传咨询. 北京：北京大学医学出版社. 2007
5. Jain KK. Applications of biochips: from diagnostics to personalized medicine. Curr Opin Drug Discov Devel, 2004, 7: 285-289
6. Fischer A, Cavazzana-Calvo M. Gene therapy of inherited diseases. Lancet. 2008, 371: 2044-2047
7. Hohmann EL. Gene therapy-still a work in clinical and regulatory progress. N Engl J Med, 2009, 361: 193-195

第十三章　表观遗传学

一个多细胞生物机体的不同类型细胞的基因型是完全一样的，然而它们的表型是各不相同的，这是由于不同类型的细胞之间存在着基因表达模式(gene expression pattern)的差异。也就是说，决定细胞类型的不是基因本身，而是基因表达模式。通过细胞分裂来传递和稳定地维持具有组织和细胞特异性的基因表达模式对于整个机体的结构和功能协调是至关重要的。

基因表达模式在细胞世代之间的可遗传性并不依赖细胞内 DNA 的序列信息。基因表达模式的信息标记，或者称之为表观遗传修饰(epigenetic modification)主要有两类，一是 DNA 分子的特定碱基的结构修饰(如胞嘧啶的甲基化)；二是由于组蛋白修饰引起的染色质构型重塑(chromatin remodeling)。

通过有丝分裂或减数分裂来传递非 DNA 序列信息的现象称为表观遗传(epigenetic inheritance)。表观遗传学(epigenetics)则是研究不涉及 DNA 序列改变的基因表达和调控的可遗传的变化，或者说是研究从基因演绎为表型的过程及其机制的一门新兴的遗传学分支。表观遗传的异常会引起表型的改变，机体结构和功能的异常，甚至导致疾病的发生。表观遗传学正在成为医学遗传学的一个重要组成部分。

第一节　表观遗传修饰机制

一、DNA 甲基化和组蛋白修饰

DNA 甲基化(DNA methylation)主要表现为基因组 DNA 上的胞嘧啶第 5 位碳原子和甲基间的共价结合，胞嘧啶由此被修饰为 5-甲基胞嘧啶(5-methylcytosine,5-mC)。哺乳动物基因组 DNA 中 5-mC 约占胞嘧啶总量的 2%～7%，绝大多数 5-mC 存在于 CpG 二联核苷酸(CpG doublets)。哺乳类动物基因组中的 CpG 二联核苷酸出现的频率远低于 4 种碱基随机排列所预期的频率，但对蛋白质编码基因

而言，CpG 二联核苷酸并不呈现基因组总 DNA 中的低频率。在结构基因的调控区段，CpG 二联核苷酸常常以成簇串联的形式排列。结构基因 5′端附近富含 CpG 二联核苷酸的区域称为 CpG 岛（CpG islands）。

在哺乳类基因启动子中，约 40% 含有 CpG 岛。CpG 岛中的 5-mC 会阻碍转录因子复合体与 DNA 的结合，所以 DNA 甲基化一般与基因沉默（gene silence）相关联；而非甲基化（non-methylated）一般与基因的活化（gene activation）相关联。去甲基化（demethylation）则往往与一个沉默基因的重新激活（reactivation）相关联。

基因的甲基化型（methylation pattern）通过 DNA 甲基转移酶（DNA methyltransferases，DNMTs）来维持。DNMT 将 S-腺苷甲硫氨酸（S-adenosylmethionine，SAM）上的甲基转移至胞嘧啶核苷酸的 5 位碳原子（图 13-1A）。在哺乳动物细胞中，已经发现了三种具有催化活性的 DNMTs，即 Dnmt1，Dnmt3a 和 Dnmt3b。当一个甲基化的 DNA 序列复制时，新合成的 DNA 双链呈半甲基化（hemimethylated），即只有母链有完整的甲基化标记，而另一条链会经 Dnmt1 的催化在与母链上 5-mC 对称的位置上使相应的胞嘧啶甲基化（图 13-1B）。因此，Dnmt1 主要在 DNA 复制中维持 DNA 甲基化型的存在，Dnmt3a 和 Dnmt3b 则是不依赖半甲基化 DNA 分子中的甲基化模板链而从头开始合成 5-mC 的从头甲基化酶（de novo methylase）。这些 DNMTs 以及 DNA 去甲基化酶（DNA demethylase）在 DNA 甲基化型的建立、维持和改变中相互协调，是基因表达表观遗传学调节的重要基础之一。

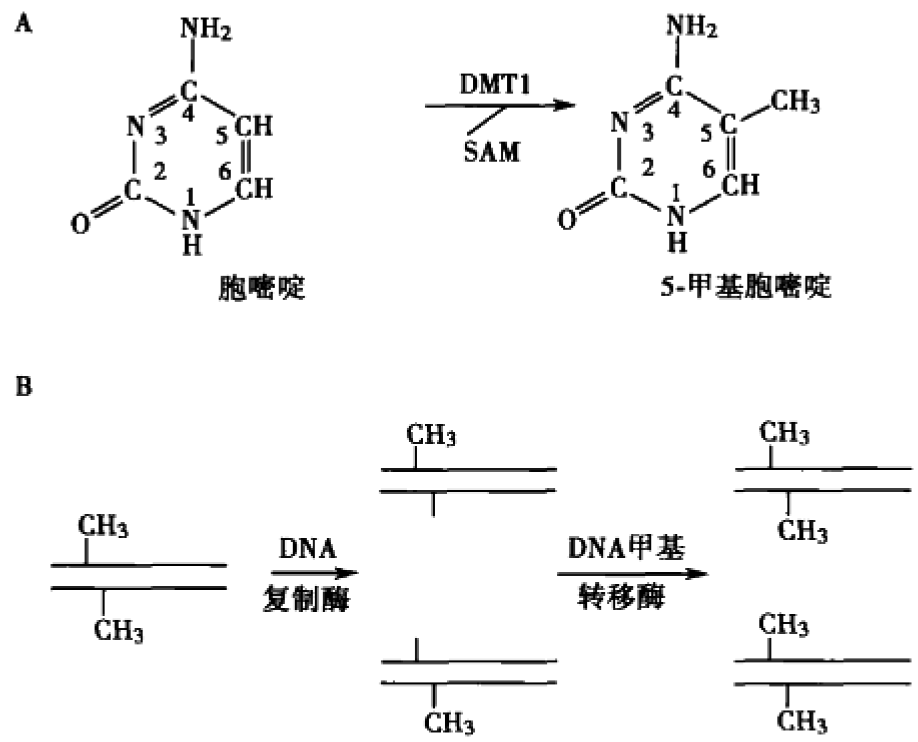


图 13-1 胞嘧啶甲基化及甲基化型的维持机制

A. 胞嘧啶甲基化反应；B. DNA 复制后甲基化型的维持

构成核小体的组蛋白氨基端可以被多种酶进行各种修饰，如磷酸化、乙酰化、甲基化和泛素化等，组蛋白的这类修饰可以改变 DNA-组蛋白的相互作用，使染色质的构型发生改变，称为染色质构型重塑。组蛋白中不同氨基酸残基的乙酰化通常预示着开放的常染色质（euchromatin）构型以及转录活性区域；而组蛋白的甲基化既与浓缩的异染色质（heterochromatin）以及基因转录受抑相关，也可以出现在转录活性区域。组蛋白的修饰可以相互影响，并和 DNA 甲基化相互作用。组蛋白氨基端的大量修饰形成不同的组合，构成了可被转录复合物识别的组蛋白密码（histone code）。例如，组蛋白 H3 的第 9 位赖氨酸残基（H3K9）在组蛋白乙酰化转移酶（histone acetytransferase，HAT）作用下的乙酰化修饰与基因活性表达相关联，一旦经组蛋白脱乙酰酶（histone deacetylase，HDAC）催化脱去乙酰基，再经组蛋白甲基转移酶（histone methyltransferase）作用在同一位置加上甲基，就会形成一个异染色质蛋白（heterochromatin protein 1，HP1）或其他抑制性染色质因子的结合位点。HP1 的结合会导致 DNA 分子

上特定 CpG 岛的甲基化和稳定的基因沉默。但是，组蛋白 H3 第 4 位赖氨酸残基(H3K4)或第 36 位赖氨酸残基(H3K36)的甲基化修饰则与基因转录激活相关联。所以，组蛋白的修饰对基因表达的影响展示了生物系统的复杂性；染色质蛋白也并非只是一种包装蛋白，而是在 DNA 和细胞其他组分之间构筑了一个动态的功能界面。

二、非编码小 RNA 分子的调节

近年来，在真核细胞中发现大量转录的一类 RNA 分子，这些 RNA 分子既不行使 mRNA 的功能，也缺乏 tRNA 和 rRNA 的功能，被称之为非编码 RNA(non-coding RNAs，ncRNAs)。其中，大约由 20～30 核苷酸(nt)构成的小 ncRNAs 分子已被公认为在调节真核基因组表达和功能中起关键作用。根据小 ncRNAs 分子的起源、结构、所结合的效应子蛋白以及功能作用，可以分为三类：short interfering RNAs(siRNAs)；microRNAs(miRNAs)和 piwi-interacting RNAs(piRNAs)。siRNAs 和 miRNAs 的前体是双链分子，其中 siRNA 前体是较长且完全配对的双链分子，miRNA 的前体分子呈现为不完全配对的发夹结构；而 piRNAs 的前体则是单链分子。siRNAs 和 miRNAs 主要和 Argonaute 蛋白家族 Ago 成员结合，而 piRNAs 则与 Piwi 成员结合。miRNAs 主要表现为内源性基因的调节子，siRNAs 则表现为维护基因组完整性的防卫者，主要针对外源或入侵的核酸起作用，如病毒，转座子和转基因等。piRNAs 主要存在动物细胞中，并集中在生殖细胞中行使功能。

1993 年，Ambros 和同事在线虫(*Caenorhabditis elegans*)中首次发现了 miRNA 分子 *lin*-4，其功能是控制发育时程基因的内源性调节。五年后，Fire、Mello 和同事报道了外源性双链 RNA(dsRNA)可通过 RNA 干扰(RNAi)机制特异性的沉默基因。1999 年，在植物中发现基因的沉默伴随着约 20～25nt RNAs 的出现。此后进一步发现，dsRNAs 可以直接转换成 20～23 nt siRNAs。

Dicer 酶是特异识别 dsRNA 的核糖核酸酶Ⅲ(Ribonuclease Ⅲ，RNaseⅢ)家族成员之一，它能以 ATP 依赖方式将内源或外源 dsRNA 前体分子加工成典型的约 21nt 长的双链分子。然后这个双链产物经过解链，其中一条链作为指导链与效应子 Ago 蛋白稳定结合，形成不同种类的 RNA 诱导沉默复合体(RNA-induced silencing complex，RISC)，另一条链(过客链)则被丢失。根据 Watson-Crick 碱基配对原理，指导链识别靶向 RNA 分子，最终，RISC 可以分别通过抑制转录或翻译、促进异染色质形成以及加速 RNA 或 DNA 降解等机制，从而实现对各种靶基因的表达调控(图 13-2)。

迄今已知 miRNA 调节大约 30% 人类基因的表达。miRNA 可以通过靶向 DNA 或组蛋白修饰酶等表观遗传复合物而实现调节作用。例如，最初在肺癌细胞株中发现，一个包括 miR-29a、miR-29b 和 miR-29c 的 miRNA 家族(miR-29s)可以直接调节 DMNT3a 和 DMNT3b；当重新表达 miR-29s，可破坏从头 DNA 甲基化，从而导致肿瘤细胞的 DNA 普遍低甲基化。在肺癌细胞中，肿瘤抑制基因(tumor suppressor genes，TSGs)是由于启动子超甲基化而被表观沉默的。miR-29s 可使 TSGs 启动子的 CpG 岛去甲基化使其从新表达，从而诱导肿瘤细胞凋亡和生长抑制。这些发现揭示：miRNAs 可以通过直接调节表观遗传过程而间接调节基因的表达。在急性髓系白血病细胞中，miR-29b 除了直接作用于 *DMNT3a* 和 *DMNT3b* 外，还可以通过直接作用于 *DNMT1* 基因的转录因子 SP1，从而间接沉默该基因。此外，miR-148a 和 miR-148b 也参与调控 *DNMT3b* 的表达：通过结合到 *DNMT3b* 基因 mRNA 编码区域的一个特异位点，miR-148 家族可以调节这个从头开始合成的 DNMT，并与 *DNMT3b* 几个不同剪切变异子的形成密切相关。另一方面，由于 miR-148a 本身也受表观遗传的调控，其启动子在不同的肿瘤组织中表现为超甲基化状态。这些证据说明，靶向表观遗传复合物的 miRNA 还存在着自我扩增的表观遗传环路。

miRNAs 也调节 HDACs 以及多梳抑制复合物(polycomb repressive complex，PRC)基因的表达。*HADC4* 是 miR-1 和 miR-140 调节的直接靶点，而 miR-449a 可与 *HDAC1* 的 3′-UTR 区域结合。*HADC1* 在几类肿瘤细胞中被上调，miR-449a 在前列腺癌细胞中重新表达，可使 HDAC1 水平降低，诱导细胞周期阻滞，细胞凋亡以及衰老表型。EZH2 是 PRC2 的催化亚基，通过使组蛋白 H3 第 27

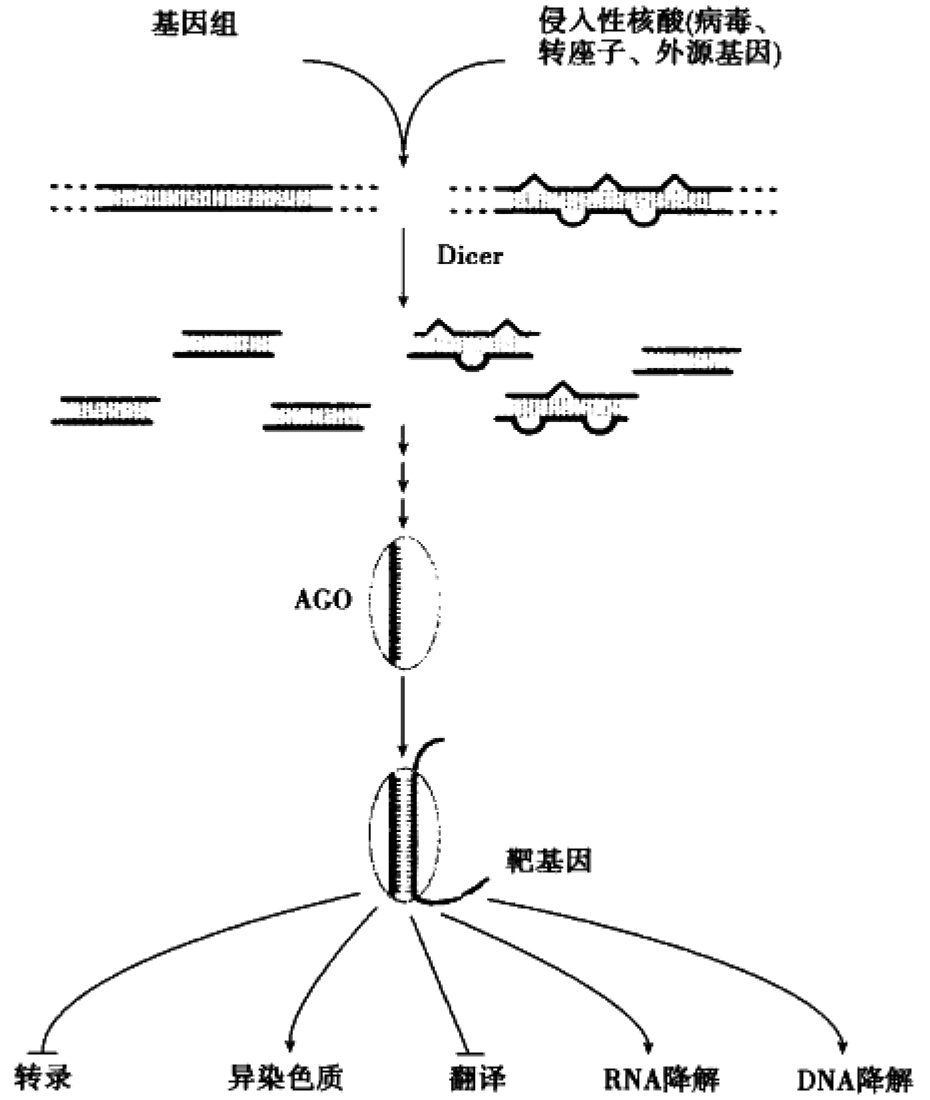


图 13-2　miRNA 和 siRNA 来源及功能途径
（引自 Carthew RW,Sontheimer EJ,2009）

位赖氨酸残基三甲基化（H3K27me3）可促使异染色质形成，导致几个 TSGs 沉默。在前列腺癌细胞株和原发肿瘤组织中，miR-101 在肿瘤发展中表达下调，与 EZH2 的增加逆向相关。进一步证据表明 miR-101 的确可以在前列腺和膀胱癌模型中直接靶向 *EZH2*。miR-101 介导的 *EZH2* 抑制阻止了肿瘤细胞的增殖和克隆形成。也揭示 miR-101 通过对肿瘤表观基因组的调节而具有的肿瘤抑制作用。

miRNA 对表观遗传的调节存在细胞或种间特异性。例如，在小鼠胚胎干细胞中，miR-290 簇直接作用于 *DNMT3* 基因的抑制子 RBL2。在 Dicer 缺失的胚胎干细胞中，miR-290 簇不表达，而 RBL2 却过表达，从而导致端粒重组和异常的端粒延长。若 miR-290 簇重新表达，则可以逆转这种现象。然而，在 Dicer 被抑制的人胚肾（HEK293）细胞中，miR-290 簇缺乏对从头 DNMTs 的调节作用。

siRNA 通过诱导异染色质的形成，也可以实现对基因表达的调控。在裂殖酵母 *Schizosaccharomyce pombe* 中，含 Ago1 的效应子构成了 RNA 诱导转录沉默（RNA-induced transcription silencing，RITS）复合物，通过结合的 siRNAs 可被导向特异的染色体位点，例如，着丝点重复序列。RITS 复合物与 RNA 多聚酶 Ⅱ 的直接相互作用可加速新生转录物对 siRNA 的识别。在组蛋白甲基转移酶（HMTs）介导下，RITS 复合物的结合可促进组蛋白 H3 第 9 位氨基酸残基赖氨酸的甲基化（H3K9），进而诱导募集具有染色质结构域的 Swi6 蛋白（HP1 在酵母中的同源蛋白），最终导致染色质浓缩。RITS 对新生转录物结合也可激活 RNA 依赖的 RNA 聚合酶复合物（RDRC），RDRC 利用其 RdRP 亚基（Rdp1）产生次级 siRNAs，进而加强和扩散沉默效应。

piRNA 以限定模式在哺乳类睾丸中表达，主要表现为在配子形成过程中对转座子元件的沉默作用，是生殖细胞发育所必须的。2008 年，K. Miyagawa 等人证实：缺失 piRNA 相互作用蛋白 MIL1 和

MIWI2 可导致雄性生殖细胞逆转座子从头 DNA 甲基化丧失。提示在哺乳动物生殖细胞中存在着 RNA 指导的 DNA 甲基化机制，为基因表达的表观遗传调节通路提供了新的证据。这一系列实验揭示：真核细胞中存在着一个由小 RNA 分子调节，组蛋白结构修饰和 DNA 甲基化系统组成的一个表观遗传修饰网络，能动地调控着具有组织和细胞特异性的基因表达模式。机体的表观遗传模式的变化在整个发育过程中是高度有序并严格受控的。

三、基因组印迹

基因组印迹（genomic imprinting）是表观遗传调节的一种形式，是指两个亲本等位基因的差异性甲基化造成了一个亲本等位基因的沉默，另一个亲本等位基因保持单等位基因活性（monoallelic activity）。1956 年 A. Prader 和 H. Willi 等医师报道了一种因父源染色体 15q11-q13 区段缺失而引起的一种儿童早期发育畸形，特征是肥胖，矮小，并伴有中度智力低下，称为 Prader-Willi 综合征（Prader-Willi Syndrome，PWS）。1968 年 H. Angelman 医师又报道了因母源染色体同一区段缺失引起的一种在儿童期以共济失调，智力严重低下和失语等为特征的综合征，称为 Angelman 综合征（Angelman syndrome，AS）。PWS 和 AS 这一对综合征表明父亲和母亲的基因组在个体发育中有着不同的影响，这种现象是典型的基因组印迹。在有些 PWS 和 AS 患者中也观察到了多种该区段的微小染色体缺失，通过对小缺失的分析发现这段缺失集中的区域有成簇排列的富含 CpG 岛的基因表达调控元件，称为印迹中心（imprinting centers，ICs），也称为印迹控制区（imprinting control regions，ICRs）或印迹控制元件（imprinting control elements，ICEs）。在父源和母源染色体上，这些调控元件的 CpG 岛呈现甲基化型的明显差异，即父源和母源染色体上的 ICs 的甲基化呈现出分化状态，或者叫差异甲基化（differential methylation）。例如，在 15q11-q13 区有一段定名为 SNRPN 的长度为 430bp 的调整区段，它含有 23 个 CpG 二联核苷酸。在遗传自母源的染色体上的 23 个 CpG 二联核苷酸完全被甲基化，而遗传自父源的染色体的则全都为非甲基化。实验进一步表明，这种呈差异甲基化的 ICs 也是该区段的邻接基因的表达调控元件。1997 年 T. Kishino 等的研究证实位于该区段的泛素-蛋白连接酶（ubiquitin-protein ligase）的编码基因 UBE3A 突变或失表达（loss of expression）可以引起 AS。AS 患者中该基因的失表达是由于母源染色体上包括 ICs 在内的染色体片段缺失所致，而 PWS 患者是由于该区段的多个父源印迹基因的错误表达（misexpression）所致。这种错误表达的后果是邻近基因启动子的从头甲基化和随后产生的基因沉默。

Beckwith-Wiedemann 综合征（Beckwith-Wiedemann syndrome，BWS）是一种过度生长综合征（overgrowth syndrome）常伴有肥胖和先天性脐疝等症状，并有儿童期肿瘤易患倾向。它起源于染色体 11p15.5 区段的多种能造成该区段印迹基因表达失衡的遗传学和表观遗传学调节机制异常。在该区段的一个长约 1Mb（相当于 1000kb）的片段中至少有 12 个成簇排列的印迹基因（imprinted genes），其中有些呈父源等位基因表达模式，另一些呈母源等位基因表达模式，这些基因分属两个印迹域（imprinted domain），它们的印迹状态分别受控于两个 ICs。在第一个 IC 中，主要有胰岛素样生长因子（insulin-like growth factor 2，*IGF2*）基因，*H19* 基因，和一个富含 CpG 岛的差异甲基化区域（differentially methylated region，DMR）三者的排列次序是：5′-*IGF2*-DMR-*H19*-3′。*IGF2* 是一种父源等位基因表达的胚胎生长因子，它的表达上调对 BWS 的病理过程非常重要。*H19* 是一种母源等位基因表达的 pol Ⅱ 转录子，它的转移物是丰度很高但功能不详的非编码 RNA，并不翻译为蛋白质。DMR 是一个印迹调控区，它借助差异甲基化，以及它特有的染色体屏障调节蛋白 CTCF 结合位点，对 *IGF2* 和 *H19* 进行交互式的印迹调节（reciprocal imprinting regulation）（图 13-3）。*H19* 和 *IGF2* 的表达要竞争位于 *H19* 3′ 下游的一个增强子。在母源染色体上，DMR1 是非甲基化的，它允许锌指蛋白 CTCF 与它相结合，从而隔断了 *IGF2* 和位于 *H19* 下游的增强子，所以该增强子只活化 *H19* 的转录。在父源染色体上，DMR1 是甲基化的，它不仅使 *H19* 基因沉默，CTCF 也因此不能与之结合，结果是父源 *IGF2* 基因在增强子作用下活化表达。在这个印迹调控区，相对增强子作用而言，DMR1 起了一个染色质屏障作用，

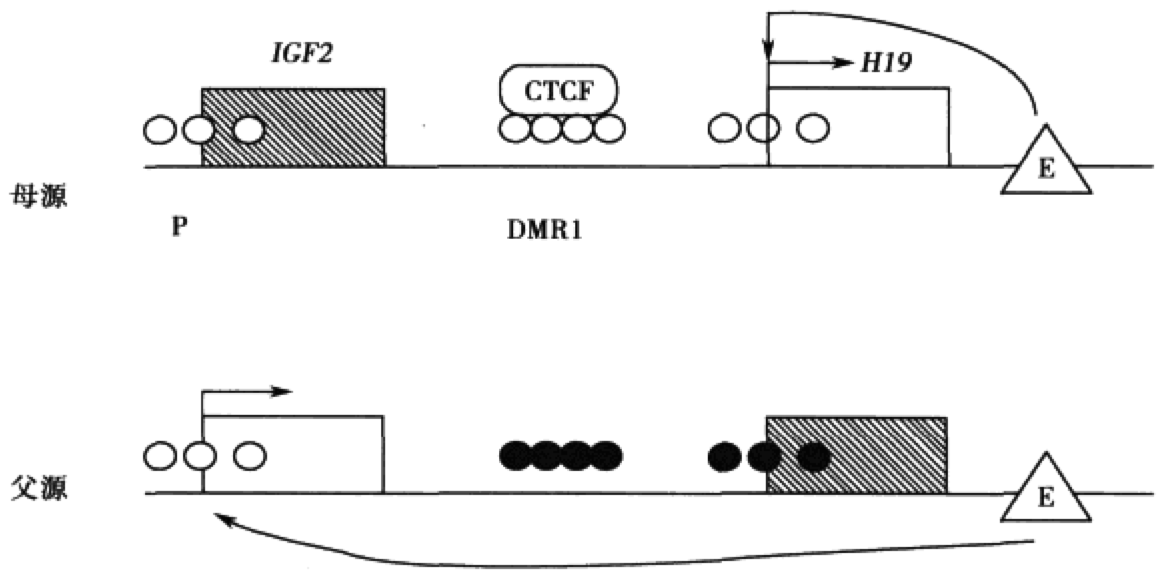


图 13-3　启动子(P)、差异甲基化区(DMR1)、锌指蛋白(CTCF)和增强子(E)对 *IGF2* 和 *H19* 的交互易换式印迹调节模式示意图
○ 为非甲基化 CpG 岛　● 为甲基化 CpG 岛

被称为隔离子(insulator),在印迹中起关键调控作用。此外,该区段的第 2 个印迹调控区也对包括编码细胞周期素依赖的激酶抑制蛋白的基因 *CDKNIC*($p57^{KIP2}$)在内的多个与细胞分裂周期相关的基因进行类似的调节。这个印迹调控区内基因的印迹失调会导致细胞的恶性生长。此外,比较基因组学分析表明,在人染色体 14q32 区也有一个与 11p15 区的 *IGF2/H19* 印迹域非常类似的印迹基因 *DLK1/GTL2* 印迹域。*DLK1* 编码一个含 6 个表皮生长因子重复基序(motif)的跨膜蛋白,也呈父源等位基因表达模式,位于 *DLK1* 下游的 *GTL2* 也编码不被翻译的 RNA,两者之间也有 CTCF 特异结合位点。所以 BWS 提供了一个具有一定典型意义的研究印迹机制的模型,尽管印迹的机制还有多种模式,如正义和反义 RNA 竞争模式,启动子特异性的交互印迹模式和双印迹中心模式等。但 *IGF2/H19* 模式或者说增强子/染色体屏障调控模式的确有助于了解基因表达的协调的机制,及其在生长发育和抑制肿瘤发生中的重要作用。

迄今已发现的印迹基因约有 80 个,大多成簇排列。虽多数印迹基因的作用机制尚不清楚,然而几乎都与 DNA 甲基化型的异常相关联。值得注意的是涉及不同亲本来源的印迹基因的 DNA 甲基化型都是在生殖细胞成熟过程中建立的(图 13-4)。也就是说,基因组印迹是性细胞系的一种表观遗传修饰,这种修饰由一整套分布于染色体不同部位的印迹中心来协调,印迹中心直接介导了印迹标记的建立及其在发育全过程中的维持和传递,并导致以亲本来源特异性方式优先表达两个亲本等位基因中的一个,而使另一个沉默。研究表明,在哺乳动物中相当数量的印

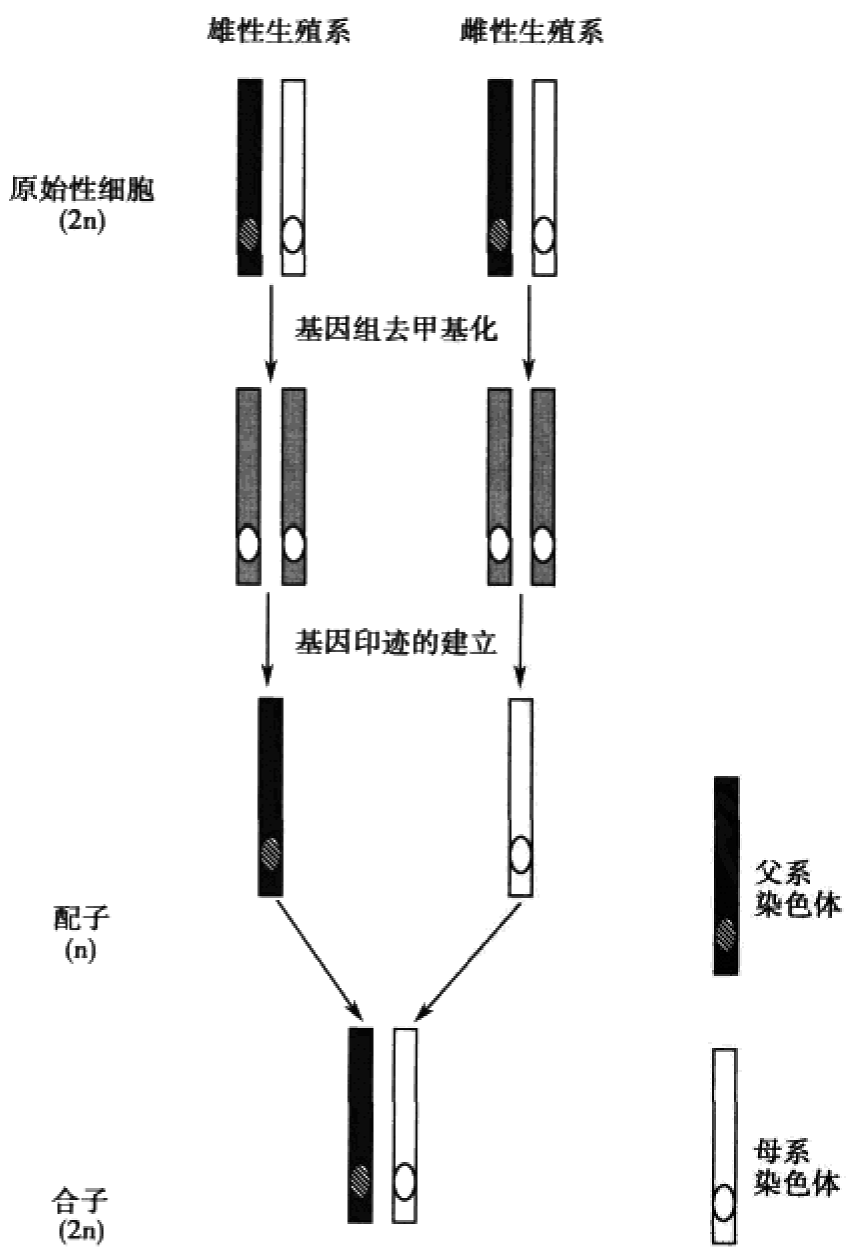


图 13-4　亲代基因组印迹在生殖系的重新编程

迹基因是与胎儿的生长发育和胎盘的功能密切相关的。这对于胚胎发育中胚胎和胎盘组织的基因表达调控非常关键。哺乳动物孤雌生殖的不可能,以及通过哺乳动物体细胞核移植来克隆动物的实验频频失败的原因之一,很可能是缺乏来自精子和卵细胞的大量印迹基因之间的协调表达。

四、基因表达的重新编程

发育是一个高度有序的生物学过程,是从一个全能的受精卵开始到建成一个由200多种具有组织和细胞特异性的、结构和功能各异的细胞组成的整体的过程。组成机体的各个组分互相协同能执行精细、复杂且相互协调的功能,如物质和能量代谢,对病原生物的抵御和免疫能力的获得,高级神经系统功能网络的建立,两性生殖细胞的发生、成熟和受精后的新生命孕育,以及与复杂多变环境之间的相互作用等。对于一个生物机体来讲,所有结构和功能各不相同的细胞虽具有完全一样的基因组,却有着很不一样的基因表达模式。与建立和维持组织和细胞特异性基因表达模式的相关细胞信息,必须是可以通过细胞分裂而遗传的,同时也应该具备被删除和重建的可能性。表观遗传修饰,如DNA甲基化和与DNA相结合的蛋白质复合物,对于稳定且可遗传的染色质构型的维持和基因表达的调控起着重要的作用。

1997年I. Wilmut和K. H. Campbell等通过一只6岁的成年母羊的乳腺上皮细胞核的移植实验,经过277次实验,终于获得了哺乳动物体细胞克隆的第一次成功,迎来了名为“多莉”(Dolly)的克隆绵羊的诞生。2006年,S. Yamanaka小组成功利用四个转录因子(Oct4,Sox2,Klf4和cMyc)从成熟的体细胞制备成诱导型多潜能干细胞(induced pluipotent stem cells,iPSCs)。这些事实雄辩地证明:在哺乳动物中,一个来自高度分化的体细胞仍然保持发育成为完整个体的能力;也就是说在分化过程中,施加在基因组的发育限制并不是永久的遗传改变,而是可逆的表观遗传修饰。哺乳动物细胞的分化是通过基因表达水平的一系列有序演化,以及细胞核和细胞质内环境的相互作用来实现的。

在自然条件下,早期原始生殖细胞(primordial germ cells,PGC)携有体细胞样的表观遗传型,在PGC进入性腺前后,原有的表观基因组开始被删除,随之在两性生殖细胞中建立性别特异性和序列特异性的表观遗传型。在受精过程中,精子进入成熟的卵细胞后,精卵融合形成的受精卵基因组在卵细胞质的生理环境中,会启动与胚胎发育相关且有严格时空特异性的基因表达程序,即删除在生殖细胞成熟过程中建立的,除印迹基因以外的全部表观遗传修饰标记,重新建立胚胎发育特有的表观基因组(epigenome)(图13-5)。也就是通过系统重建表观遗传修饰为胚胎发育中的基因表达重新编程(reprogramming)。只有经过重新编程的表观基因组才具有发育的全能性,满足胚胎所有细胞发育和专一性分化的需要,才能为胚胎发育和分化发出正确的指令,小鼠胚胎的重新编程在着床前就完成了。胚胎发育中表观基因组重新编程的误差将会导致多种表观遗传缺陷性疾病。然而克隆动物的表观基因组更接近于来自成年动物的供核细胞,这很可能是体细胞核移植克隆实验成功率极低的主要原因,也就是说体细胞核的重新编程往往难以成功。

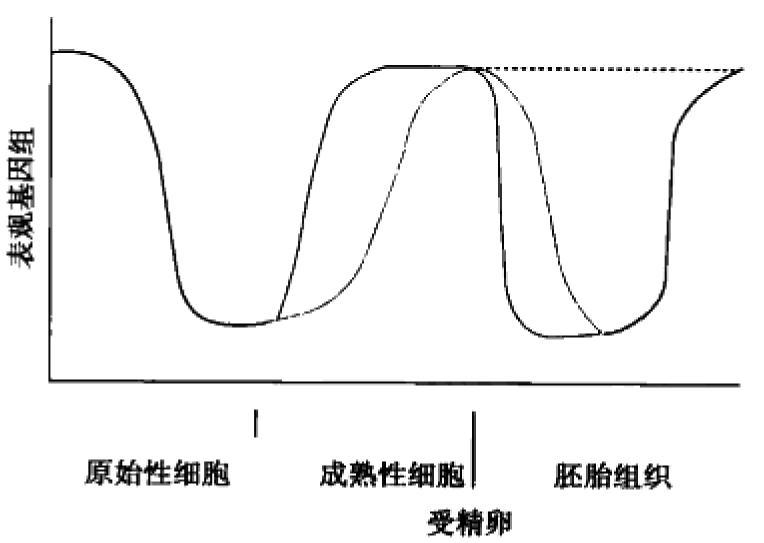


图13-5 个体发育过程中表观基因组的重新编程
早期原始生殖细胞在沿着生殖系统管腔移行时,原属体细胞型的表观遗传修饰(包括基因组印迹)会被删除。在生殖细胞发生与成熟过程中表观遗传标记重新建立(蓝线表示精子分化,红线表示卵细胞分化)。受精后会进行除印迹基因(由黑色虚线表示)以外的表观遗传修饰的删除与重建,重建后的表观基因组在组织特异性定型后被稳定地维持

另一个值得注意的问题是表观遗传修饰的重新编程对环境变化非常敏感。例如,在动物实验中,改变胚胎培养液不但会引起异常甲基化和印迹基因*IGF2*和*H19*的表达失调,甚至造成印迹性疾病。基于此观点有人检查了人工辅

助生育后代的情况。因为辅助生育是在配子生成和胚胎发育早期干预了生殖，而这个时期正是表观遗传编程获得和维持的关键时期。K. H. Orstavik 等曾报道经卵细胞胞质精子注射(intracytoplasmic sperm injection,ICSI)辅助后出生的儿童中，存在 PWS/AS 和 BWS 发生率增高的现象，并在患儿中检测到包括 *H19*,*IGF2* 在内的多个印迹基因表达异常。这些结果提示有必要对经辅助生育技术孕育的孩子进行表观遗传学监测。

1992 年 Q. Jaenish 运用基因工程小鼠较为全面地探讨了重新编程问题。他发现如果敲除小鼠的 DNA 甲基转移酶基因，突变小鼠的胚胎会出现多个器官的一系列异常表型，并都在胚胎发育早期夭折。实验有力地证明甲基化对于胚胎存活的重要性。

五、X 染色体失活

1961 年 M. F. Lyon 就提出了关于雌性哺乳动物体细胞的两条 X 染色体中会有一条发生随机失活的假说，并认为这是一种基因剂量补偿的机制。但 X 失活的机制一直是不清楚的。X 失活是典型的表观遗传现象，而且是以整条染色体为靶标的表观遗传修饰的一个特例。

1996 年 G. D. Penny 等发现 X 染色体的 Xq13.3 区段有一个 X 失活中心(X-inaction center,Xic),X 染色体失活从 Xic 区段开始启动，然后扩展到整条染色体，Xic 长约 1Mb，包括 4 个已知基因：*Xist*、*Xce*、*Tsix* 和 *DXPas34*。X 染色体失活特异性转录因子(X-inactive specific transcript, *Xist*)基因，是 X 染色体上启动转录最早的基因，但它的转录产物没有开放阅读框(ORF)。两条 X 染色体的 *Xist* 基因都能从上游启动子启动 *Xist* RNA 的稳定转录，但随后只有一条 X 染色体产生的 *Xist* RNA 将这条染色体自身整体包裹，并启动异染色质化和失活过程，而另一条 X 染色体转录的 *Xist* RNA 会很快裂解，这条 X 染色质则呈常染色质状态，整条染色体上的基因都具有表达活性。值得注意的是 *Xist* RNA 在失活的 X 染色体表面呈现锚钉样排列，提示它可能与染色体上特定的蛋白质相结合而形成稳定的结构(图 13-6)。

Xce(X-chromosome controlling element)基因主要影响 X 染色体随机失活的选择，当 *Xce* 处于纯合状态时，在体细胞中的 X 染色体的失活是完全随机的，而在杂合状态时，失活就不是完全随机的。*Tsix* 基因是位于 *Xist* 下游的顺式调控元件，其中包含 CTCF 的结合位点，提示 CTCF 与 *Tsix* 可能协同起着 *Xist* 外源开关的功能。*DXPas34* 基因富含 CpG，包括一个 15kb 的微卫星重复序列，提示对 X 染色体失活有一定调控作用。失活 X 染色体有两个显著特点，一是组蛋白 H4 不被乙酰化；二是 CpG 岛的高度甲基化。

雌性体细胞 X 染色体的随机失活始于对 X 染色体的计数和对未来活性或失活 X 染色体的选择，并以即将失活 X 染色体的 *Xist* 基因转录上调为标志。任何 X 染色体被失活的概率随着 X 染色体对常染色体的比率增加而增加，提示 X 染色体编码的激活子参与了 X 染色体随机失活的计数过程。为了启动 X 染色体失活，*Xist* 必须超越 *Tsix* 参与设定的阈值。在抵消 *Tsix* 过程中，X 染色体编码的激活子对 *Xist* 表达具有剂量依赖性激活作用；常染色体编码的抑制子则表现为对 *Xist* 的剂量依赖性抑制作用。已有证据表明：在细胞核内，X 染色体编码的激活子浓度确实可以触发 X 染色体失活。2009 年，J. Gribnaul 等人发现：X 连锁基因 *Rnf12* 编码的 E3 泛素连接酶 RLIM 能依剂量依赖的特点反式激活 *Xist*。因此，雌性体细胞可以通过调节 RLIM 蛋白表达量去失活其中一条 X 染色体(图 13-6A)。*Rnf 12* 位于 *Xist* 上游 550kb 处。作为激活子，RLIM 蛋白可以通过直接激活 *Xist* 或者间接干扰 *Tsix* 顺式调节位点以及其他调控元件(如 *Xce*)而起作用。当激活作用超越抑制作用时，*Xist* 的 RNA 被转录，从而启动 X 染色体失活，因此也顺式关闭 *Rnf12* 的一个等位基因。

在小鼠胚胎干细胞分化过程中(图 13-6B)，正负信号必须在 Xist 位点被整合，只有信号强度达到刺激阈(虚线)时 *Xist* 才能表达。在未分化的细胞中，Oct4 和 Nanog 等多潜能因子可以阻止 *Xist* 的表达；在分化早期阶段，多潜能因子被下调，整合的 *Xist* 刺激逐渐增强并仅在雌性细胞中超过激活阈值；此外，与雄性体细胞相比，雌性体细胞中还包含较多的 X 染色体连锁的 *Xist* 激活子；加之在该阶段发生的同源配对(灰色箭头)打破了 *Xist* 之间的对称。因此，阻止了 *Xist* 两个等位基因同时被激活的趋

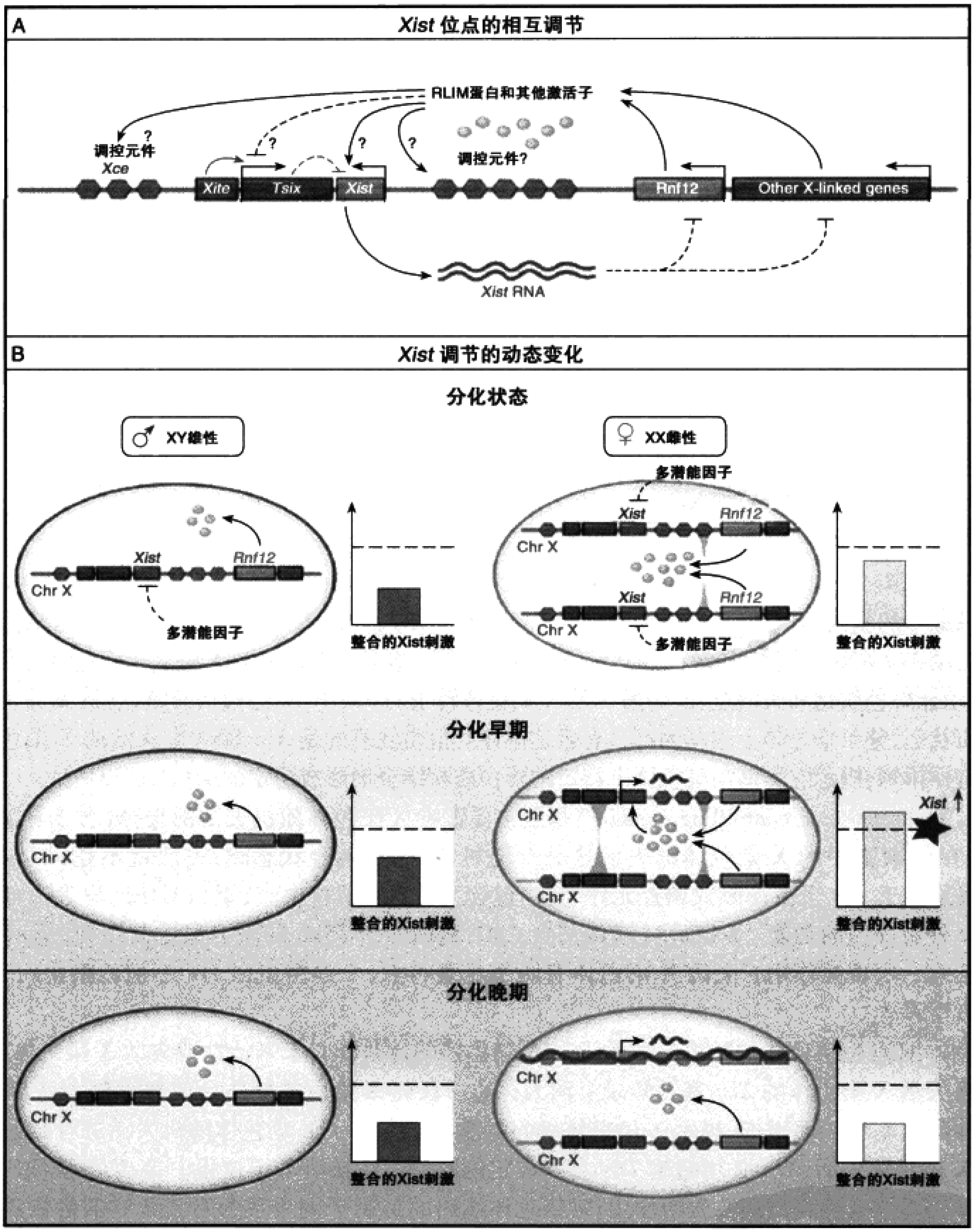


图 13-6 *Xist* 位点介导 X 染色体失活过程示意图

（引自 Nora EP，Heard E，2009）

势。在分化晚期，*Xist* RNA 介导的顺式沉默下调 X 染色体连锁的激活子，导致刺激信号减弱，最终使得 X 染色体活性在雌雄个体细胞内达到平衡。尽管 X 染色体失活的分子机制还有待于进一步阐明，但无疑会成为今后表观遗传学研究的重要内容。

第二节　表观遗传与疾病

表观遗传调控取决于多层次表观遗传修饰的互相作用。由表观遗传修饰异常引起的疾病主要可分为两大类，一类是在发育的重新编程过程中造成的特定基因表观遗传修饰的异常，有人称之为表观

突变(epimutation);另一类与表观遗传修饰的分子结构与功能相关的蛋白质编码基因有关,如 DNA 甲基转移酶基因或差异甲基化 CpG 岛结合蛋白 *CTCF* 基因的突变或表观突变。

1983 年 B. Hagberg 等报道了一种遗传性进行性神经系统疾病 Rett 综合征,患者均为女性,在出生后 7~18 个月就出现发育停滞,随后出现高级脑功能的迅速恶化和严重痴呆等症状,家系分析显示 Rett 综合征是一种 X 连锁基因突变所致。近年来的研究表明 Rett 综合征的致病基因是 X 染色体上编码 McCP2 蛋白质的基因。McCP2 是一种甲基结合蛋白(methyl-binding proteins, MBPs),能专一性地识别甲基化 CpG 岛并与之结合,其功能是作为分子榫头将染色质修饰复合物(chromatin-modifying complex, CMCs)和 DNA 甲基化区域连接在一起以阻遏基因转录。Rett 综合征患者的 *McCP2* 基因突变集中在甲基化 CpG 结合域和转录阻遏域。显而易见,这类突变会严重干扰表观遗传修饰的正常功能。值得注意的是 *McCP2* 基因的表达谱比较广泛,所以突变所造成的病理作用为什么只局限于脑内神经细胞的机制还有待于研究。

ICF 综合征(immunodeficiency, centromeric region instability and facial anomalies syndrome)是一种罕见的常染色体隐性遗传病,它是一种变异性免疫缺陷,主要病症是不同程度的免疫缺陷,并伴以面部畸形和智力低下。在分子水平上,患者有典型随体序列Ⅱ和Ⅲ的低甲基化。多个研究小组独立发现,该病是能催化 DNA 从头甲基化的 *DNMT3B* 编码基因的突变所致。患者至少有两种同型免疫球蛋白的减少或缺失,并造成细胞免疫缺陷。此外,淋巴细胞分裂中 1 号、9 号和 16 号等多条染色体的环着丝粒区域的不稳定性也明显增高。这些区域含一种通常被甲基化的卫星 DNA 序列,但在 ICF 患者中几乎完全是非甲基化的。这些序列被认为与着丝粒的功能和动粒(kinetochore)的装配有关。还有人发现 ICF 患者失活 X 染色体的 CpG 岛和两个重复序列家族 D4Z4 和 NBL2 的去甲基化。这些变化与基因组中 5-mC 水平降低和某些染色体着丝粒周围区域重复序列的低甲基化相互吻合。基因芯片的表达分析还显示,患者淋巴细胞中部分与免疫功能调节相关的基因表达水平下调,但未观察到这些基因启动子区甲基化型的变化,提示 ICF 综合征中 *DNMT3B* 基因的突变可能通过降低转甲基活性而减少了对基因转录的阻遏作用,从而间接影响了淋巴细胞部分基因的表达模式。

1981 年 D. J. Weathrall 等报道了地中海贫血和智力低下的联系,随后证实这并非患者同时患有两种疾病,而是一种 X 连锁疾病,称为 X 连锁 α-地中海贫血/智力发育迟滞综合征(X-linked alpha-thalassemia/mental retardation syndrome; ATR-X)。*ATR-X* 基因突变会引起特征性的发育异常,如严重的智力低下、面部变形、α-地中海贫血、泌尿生殖道畸形,甚至出现性反转表型。现已证实 *ATR-X* 基因编码一个含有植物的发育同源结构域(plant hemeodomain, PHD)的蛋白质,是一种转录调节因子,通过修饰染色质的局部结构来调节转录,该蛋白质的 C 末端还含有解旋酶 SNF-2 家族成员的标志性结构域。在细胞分裂间期和中期,该蛋白质定位在着丝粒附近的异染色质区。在 ATR-X 患者中发现一些高度重复序列的甲基化型发生了改变,包括编码核糖体 RNA 的 rDNA 重复序列、Y 染色体特异的卫星 DNA 和亚端粒区重复序列等。这些重复序列区域甲基化的严重减少,加上染色质重塑解旋酶 SNF-2 结构域的存在,提示 *ATR-X* 编码的蛋白质功能可能起着将 DNA 甲基化和染色质重塑这两类表观遗传修饰连接在一起的作用。一些植物和哺乳动物的实验观察也证实 SNF-2 样蛋白的突变会导致基因组甲基化的急剧丢失。

脆性 X 染色体综合征(fragile X syndrome)是一种以智力低下为主要症状的遗传性智力障碍综合征,致病基因是位于 Xq27.3 的脆性 X 智障基因(fragile X mental retardation-1, *FMR1*)。该基因最常见的突变是 5′端非翻译区中 CCG 三核苷酸重复序列的扩展。正常人的$(CCG)_n$ 重复序列为 6~50 拷贝,扩增至 52~200 拷贝时称为前突变(premutation),扩展至 200~2000 拷贝时称为全突变(full muation),这种$(CCG)_n$ 拷贝数的扩展是随着世代而不断进行的,又被称为动态突变(dynamic mutation)。分析表明$(CCG)_n$ 重复序列扩展会引起 CCG 中 CpG 二联核苷酸的甲基化,从而使 *FMR1* 基因沉默,这种沉默还涉及染色质构型的改变,而染色质的浓缩进一步使扩展的$(CCG)_n$ 重复序列的遗传

稳定性增加。

D. Wohrle 等人进行了鼠源胚胎肿瘤细胞 *FMR1* 转基因实验，进一步为 CCG 重复序列甲基化在脆性 X 综合征发生中的分子病理学作用提供了新的线索。当将携有 CCG 重复序列并甲基化的脆性 X 染色体转入肿瘤细胞后，会导致去甲基化和 *FMR1* 基因转录的重新激活，并增加 $(CCG)_n$ 重复序列的遗传不稳定性。为了证实肿瘤细胞中确实发生了去甲基化反应，利用能诱导去甲基化反应的 5-氮脱氧胞苷(5-aza-2′-deoxycytidine)处理脆性 X 细胞，结果使 *FMR1* 基因和乙酰化组蛋白 H3 和 H4 重新联结，转录也重新被激活。这表明脆性 X 综合征患者的 *FMR1* 基因沉默最初起因于延伸重复序列的甲基化，这也是研究得最清楚的一种因特定 DNA 序列表观遗传修饰异常而导致的一种疾病表型。

表观修饰遗传对于控制基因转录和染色体稳定是十分重要的。表观遗传信号甚至可以通过阻遏特定基因转录的双链 RNA 分子在细胞间的传递来影响其他细胞的基因表达。虽然对于表观遗传修饰在疾病发生中作用的认识还很不全面，但 R. Holliday 等人已经提出了表观遗传病(epigenetic diseases)的概念，其中包括多种复杂的遗传性综合征，印迹综合征，免疫性疾病和中枢神经系统发育紊乱等，还包括衰老和癌症。

第三节　表观遗传与肿瘤

早在 1979 年 R. Holliday 就曾经提出 DNA 甲基化可能在癌变过程中起着重要的作用。1983 年 A. P. Feinberg 和 B. Vogelstein 就发现癌细胞中 DNA 甲基化的总体水平低于正常细胞。并证实肿瘤细胞的低甲基化频繁发生于重复序列和在生物进化过程中引入的外来寄生性 DNA(parasitic DNA)。在正常情况下，这类 DNA 序列是被高度甲基化的，这种特征性的甲基化反映出 DNA 甲基化和 RNAi 一样源于在进化中对外来 DNA 侵袭的防御，例如细菌中的限制-修饰系统就是以 DNA 特定序列甲基化和甲基化序列专一性核酸内切酶为基础的。实验也表明人类基因组的内源性转座子(transposon)等可移动因子的激活确实是和细胞的遗传不稳定性相关联，如肺癌和结肠癌细胞 DNA 总体甲基化水平的降低就和 *KRAS* 等癌基因的激活有关。

除 DNA 甲基化的总体水平降低之外，癌细胞往往出现局部序列的高甲基化，高甲基化通常集中在启动子等基因表达调控元件附近的 CpG 岛，从 1986 年 J. B. Baylin 等最初发现降钙素(calcitonin)基因在癌细胞中呈现出正常细胞中不存在的异常甲基化起，迄今已发现 60 多个基因在癌细胞中显现异常甲基化。例如，在肾癌和视网膜母细胞瘤中 *VHL*(Von-Hippel Landau)基因的沉默和前列腺癌中与 DNA 损伤修复有关的 *GSTPI* 基因的沉默，都可以被抑制 DNA 甲基转移酶的 5-氮胞苷重新激活，表明 DNA 甲基化的异常是癌细胞中某些抑制恶性生长的基因沉默的原因。

在人类癌细胞中还观察到了多个功能相互关联的基因，如细胞周期相关基因群和 DNA 损伤修复相关基因群同时显现肿瘤特异性甲基化。实验还提示基因表达的甲基化失活似乎是癌变的极早期事件，例如在结肠癌的癌前病变组织中发现过特定基因的甲基化，并观察到甲基化程度和整个病程演进过程相关。癌细胞中大多数沉默的肿瘤抑制基因启动子区域 CpG 岛的高甲基化，也许是一个分阶段的渐进过程，先是某些特定的 CpG 岛从头甲基化，再由此扩展至更多的 CpG 岛，最终使基因进入稳定的沉默状态。有实验提示，最初的甲基化可能起源于在 DNA 自发损伤修复中产生的 5 甲基脱氧单磷酸胞嘧啶(5mdCMP)的错误掺入。在正常情况下，5mdCMP 由一种专门的酶脱氨后转变为脱氧单磷酸胸腺嘧啶(dTMP)，再经磷酸化以 dTMP 形式掺入 DNA 分子。而在癌细胞中，相当一部分酶的活性发生了改变，致使部分 5mdCMP 又以二磷酸或三磷酸核苷的形式错误掺入 DNA。有人曾发现一个长期在体外培养的中国仓鼠卵巢细胞(CHO 细胞)分离株的 5mdCMP 脱氧酶的活性明显下降，造成两个受试基因位点的自发表观突变(spontaneous epimutation)频率明显增高，H^3-标记的 5mdCMP 掺入实验证明这个 CHO 的表观突变细胞株的确掺入了甲基脱氧胞苷。

近年来的研究还揭示了表观遗传病与肿瘤相关。例如BWS患者的Wilms瘤的发病率比对照群体高出1000倍。对肺癌，神经胶质瘤(glioma)，乳腺癌和结肠癌的分析表明*IGF2*等基因的印迹丢失(loss of imprinting,LOI)是肿瘤危险因子，也是最常见的表观遗传改变。LOI的机制还涉及CTCF和另一种印迹调控蛋白(brother of regulator of imprinted sites,BORIS)在染色体上结合靶点的甲基化状态的改变，以及印迹调控蛋白质复合体对染色质结构重塑的影响。

肿瘤形成中表观遗传修饰的病理作用并没有否定从R. Weinberg 1981年在人类膀胱上皮癌中发现*RAS*基因突变以来有关肿瘤起源的突变研究，而是为我们提供了一种不涉及DNA序列改变的病因研究途径，也为我们提供了一种新的肿瘤治疗手段，对于医师和药物研发企业来讲，它比基因治疗具有更大的吸引力。例如，2004年和2006年，美国食品药品管理署(FDA)分别批准两种去甲基化制剂用于骨髓异常增生症的临床治疗，它们能重新激活因甲基化而沉默的基因。2006年，FDA又批准了具有增强乙酰化作用的制剂，主要用来治疗皮肤T细胞淋巴瘤。

第四节 表观遗传与衰老

分化细胞的稳定性是高等生物的基本特征之一，无论是神经元这类特化的分裂后细胞(post-mitotic cells)，还是成纤维细胞或成骨细胞这样处于不断分裂的细胞(dividing cells)，都具有稳定的特征性表型。然而，在衰老的过程中某些细胞会发生年龄相关的变化，例如某个CpG岛的从头甲基化会关闭一个基因，丧失与这个基因相关的生理功能；同样，甲基化的丢失也会激活正常情况下沉默的基因，造成不恰当的异位表达(ectopic expression)。虽然在一个组织中发生异常甲基化的细胞只占少数或极少数，但却能使组织或器官呈现出表观遗传上的异质性和镶嵌性，这种在衰老过程中获得的表观遗传镶嵌性正是许多年龄相关的局灶性疾病的一个重要病因。

动脉粥样硬化和肿瘤一样也是一种局灶性增殖疾病，有遗传学病因，也有表观遗传学病因。失控的平滑肌细胞增殖会使血管变窄，最终导致心脏缺血或脑缺血。在动脉粥样硬化患者的心肌组织，动脉粥样斑块和长期在体外培养的血管平滑肌细胞中，都曾观察到雌激素受体α基因(estrogen receptor alpha gene,ERα)的启动子区域出现年龄相关的甲基化。同样的变化会不会影响血管组织其他基因还尚待研究。然而，从理论上讲，年龄相关的表观遗传镶嵌性在血管上皮细胞和平滑肌细胞中有可能促进动脉粥样硬化的发展。

随着基因组5-mC检测技术的进步，年龄相关的获得性疾病受到启动子甲基化影响的实验证据越来越多。例如在结肠成纤维细胞中，曾观察到*ERα*、*MLH1*(DNA错配修复蛋白1)、*MYOD*(生肌性转录调节因子)、*PAX6*(发育相关的成对框基因6)、*RARβ2*(视黄素受体β2)和*IGF2*(胰岛素样生长因子2)等基因启动子甲基化和随后的基因功能下降。又如，伴有胰岛素抵抗症状的糖尿病，也是由于表观遗传异常等原因导致胰岛素受体信号传导相关的一系列基因功能下降，造成不同基因启动子甲基化发生在同一组织的不同细胞中，大大增加了局灶性疾病的异质性，也反映了老年化组织的镶嵌性。实际上，类似的分析已经成为发现疾病相关基因的一条新途径。

认识到表观基因组在发育、生长和衰老过程中存在着一个动态变化的过程，以及体细胞的表观基因组有重新编程的可能性，不仅有助于我们以新的观点来探索老年病的病理机制，发展和建立新的诊断方法和药物干预的新途径，以及更加确切地评估老年病的发病危险性，还为通过环境和生活方式的改变来延缓老年病的发生和减轻老年病的严重程度提供了理论依据。将这些概念付诸于实践之前还必须解决三个问题：一是确定表观遗传修饰与特定生理或病理指标的相关性；二是证实将这些指标作为鉴别诊断的潜在可能性和技术可行性；三是通过一定规模的流行病学调查来验证实验室内的表观遗传病理发现在人群中的真实性。

第五节 表观遗传的生物学意义

2003 年，L. Jirtle 和 R. A. Waterland 用 *Agouti* 小鼠做了一个表观遗传学的经典实验。*Agouti* 基因（*A*）编码一种旁分泌的信号分子使毛囊黑色素细胞转为合成黄色素。在鼠毛生长的中间阶段，*A* 基因一过性的短暂表达在每根鼠毛的毛尖下方形成黄色条带，使野生型 Agouti 小鼠呈现特征性的棕褐色。实验者在 *A* 基因 5′端上游插入了一个源自逆转座子（retrotransposon）的 IAP（intracisternal A particle）序列，使 *A* 基因受隐含在 IAP 中的启动子调控而持续异位表达，造成毛色变黄，插入了 IAP 的 *A* 基因称为 A^{VY}（Agouti viable yellow gene allele）。然而，IAP 启动子区域 CpG 岛的甲基化又会使有些细胞中的 A^{VY}基因表达受抑，甚至沉默。这种表观遗传差异往往发生在胚胎发育早期，所以，即使在近交系同窝仔鼠中，A^{VY}小鼠也会出现不同的表型，从以黄色为主到杂以大小不等的棕褐色斑块。在造成毛色广泛变异的同时，还造成同窝仔鼠在脂肪代谢、葡萄糖耐受和肿瘤易感性等方面的差异。

实验的对象是基因型为 *a/a* 的母鼠及其孕育的基因型为 A^{VY}/a 的仔鼠。孕鼠分为两组，实验组孕鼠除喂以标准饲料外，从受孕前两周起还增加富含甲基的叶酸、乙酰胆碱等补充饲料，而对照组孕鼠只喂饲标准饲料。结果实验组孕鼠产下的仔鼠大多数在身体的不同部位出现了大小不等的棕色斑块，甚至出现了以棕褐色为主要毛色的小鼠（图 13-7）。而对照组孕鼠的仔鼠大多数为黄色，并对肥胖、糖尿病和肿瘤易感。分析表明喂以富甲基饲料的孕鼠所产仔鼠的 IAP 所含 CpG 岛的甲基化平均水平远高于对照组，转录调控区的高甲基化使原来呈异位表达的基因趋于沉默，毛色也趋于棕褐色。当然，由于种间差异，小鼠实验不能简单地外推到人，但这并不能降低这个实验的理论价值：即诸如营养这样的环境因素并不会引起 DNA 序列的改变，却可以通过改变基因的甲基化型而改变其表观遗传型，造成明显的、可遗传的表型效应。



图 13-7 富甲基饲料对孕鼠后代表型的影响
（引自 Waterland RA，Jirtle RA，2003）

Agouti 小鼠的实验有着深刻的启示：①表观遗传修饰的环境因子敏感性也许可以解释遗传学上完全一样的个体（如双生子）在不同的环境中可以产生明显的表型差异，也提示表观遗传修饰的可遗传性在基因和环境的相互作用中起着重要的作用；②小鼠基因组转座子插入位点的异常甲基化，会引起小鼠在细胞水平上的表观遗传镶嵌性，扩大了表型变化的范围。这一点对人类来讲也有深刻意义，因为转座子这类在进化过程中由外来 DNA 演化而来的所谓“寄生因子”，占了人类基因组的 35% 以上组分。这些寄生 DNA 序列大多数是被甲基化的，被沉默的。但也有一些处于低甲基化或非甲基化状态。现已在约 4% 的人类蛋白质编码基因中发现了转座子序列，甚至还发现不少基因的转录也像 A^{VY}/IAP 的异位表达那样起始于转座子隐含的启动子区域。所有这些都在暗示哺乳动物基因组中的转座子可能赋予机体相当大的表型可变性，也就是说，每一个哺乳动物个体都可能因此成为表观遗传的镶嵌体，也因此更容易在保持基因组稳定的前提下，提高机体对环境的适应能力。这对于个体发育和物种进化都具有十分重要的生物学意义。

人类基因组和多种模式生物基因组测序计划的完成为诠释基因组功能奠定了基础，也为对基因功能表达中起着某种决定作用的表观遗传学研究开拓了广阔空间。与高度稳定的基因组相比，表观

基因组是可遗传的，在一定条件下又是可逆的，处于亚稳定或准稳定状态，在个体发育和生殖细胞形成过程中是需要经历重新编程的，即使高度分化的成年哺乳动物体细胞也有重分化或再分化的潜在可能。在表观遗传研究过程中，还形成了表观遗传修饰，表观遗传突变，表观等位基因（epialleles），表观基因组，表观基因组学（epigenomics），表观遗传病和表观基因治疗（epigenetic therapy）等一系列科学概念。这些概念和思想已经成为哺乳动物克隆技术的进步和干细胞移植技术用于临床等应用性研究的理论先导。从技术上讲，表观遗传研究促进了一大批分析和监控技术的发展，如利用亚硫酸氢盐能选择性地使胞嘧啶核苷脱氨，而不作用于5甲基胞嘧啶的性质，发展和建立了基因组DNA的5-mC测序技术和限制性标记基因组筛选技术（restriction landmark genomic scanning，RLGS），还有甲基化敏感的任意引物PCR技术（methylation sensitive arbitrarily-primed PCR，MS-AP-PCR），差异甲基化杂交（differetial methylation hybridization，DMH），以及专门分析单个DNA分子上若干个CpG岛上呈串联状时完全甲基化的“甲光”（MethyLight）技术。甲光技术的最大优势是能够以万分之一的灵敏度在大量非甲基化和部分甲基化DNA序列的背景上检测出一连串CpG岛全甲基化的DNA。值得一提的是，甲基化型分析有两大优势可能发展为理想的检测或诊断对象。第一，甲基化型既能反映有关基因功能状态及与之相关的多种疾病相关的丰富信息，又具有简单的“二元化”性质，即令甲基化为“0”，非甲基化为“1”，就可以进行数字化处理，便于开展大规模和自动化监测分析。第二，DNA分子十分稳定，有可能将它和DNA的SNP分析等置于同一个技术平台。同时它又比RNA和蛋白质更便于保存和运输，并可对已经石蜡、甲醛或乙醇预处理的样本进行分析，可以开发历史上贮备的大量病理学资源。此外，最近基于染色质免疫共沉淀（chromatin immunoprecipitation，ChIP）技术以及与基因组芯片串联使用的ChIP-on-chip技术也已开始用于染色质修饰因子和结合因子的高通量检测，发展和建立了以组蛋白结构重塑为靶标的表观遗传修饰分析系统。然而，表观遗传修饰相关酶系的发现、鉴定和功能研究仍是今后取得突破的关键。

最后必须指出，表观遗传研究丝毫没有降低遗传学或基因组学的重要性。恰恰相反，表观遗传学是在以孟德尔式遗传为理论基石的经典遗传学和分子遗传学母体中孕育而生的、专门研究基因功能实现的一种特殊机制的遗传学分子学科。

Summary

While the DNA sequence in all cells of an organism is identical, each different cell type is largely defined by gene expression/repression patterns which is attributed to epigenetic regulation. Epigenetics has now been a rapidly advancing field of biological and medical research. In this chapter we discuss the different molecular mechanisms of the epigenetic machinery including: DNA methylation, histone modifications, the involvement of small non-coding RNAs, genomic imprinting, reprogramming of gene expression and X chromosome inactivation. Further discussion involves epigenetics in relation to cancer, aging and related diseases as well as the biological implication of epigenetics in mammals.

（傅继梁　彭鲁英）

参考文献

1. Lewin B. Genes Ⅶ. New York: Oxford University Press, 2000
2. Beck S, Ole KA. Epigenome—Molecular hide and seek. Weinbeim: Wiley-Vch Verlag Gmb&Co. KGak, 2003
3. Rakyan VK, Preis J, Morgan HD, et al. The marks, mechanisms and memory of epigenetic states in mammals. Biochem. J., 2001, 356: 1-10

4. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science, 2001, 293: 1089-1093
5. Murrell A, Heeson S, Cooper WN, et al. An association between variants in the IGF2 gene ad Beckwith-Wiedmann Syndrome: interaction between genotype and epigenotype. Hum. Mol. Genet., 2004, 13: 247-255
6. Laid PW. The power and promise of DNA methylation markers. Nat. Rev. Cancer, 2003, 3: 253-266
7. Feinberg AP, Tycko B. The history of cancer epigenetics. Nat. Rev. Cancer, 2004, 4: 143-153
8. Huang C, Sloan E, Boerkoel CF. Chromatin remodeling and human disease. Curr Opin Genet Dev, 2003, 13: 246-252
9. Kelly TLJ, Trasler JM. Reproductive epigentics. Clin Gemet, 2004, 65: 247-260
10. VU TH, Hoffman AR. Comparative genomics sheds light on mechanisms of genomic imprinting. Genome Res, 2000, 10: 1660-1663
11. Feinberg AP, Oshimura M, Barrett C. Epigenetic mechanism in human disease. Cancer Res, 2002, 62: 6784-6787
12. Waterland RA, Jirtle RA. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol, 2003, 23: 5293-5300
13. Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell, 2009, 136: 642-655
14. Valeri N, Vannini I, Fanini F, et al. Epigenetics, miRNAs, and human cancer: a new chapter in human gene regulation. Mamm. Genome, 2009, 20: 573-580
15. Hochedlinger K, Plath K. Epigenetic reprogramming and induced pluripotency. Development, 2009, 136: 509-523
16. Nora EP, Heard E. X chromosome inactivation: when dosage counts. Cell, 2009, 139: 865-867

第十四章 遗传服务的伦理问题

第一节 遗 传 服 务

一、遗传服务和它的历史发展

遗传服务(genetic service) 是一种医学服务,是将遗传学的知识和技术应用于解决人类自身的生、老、病、死的问题。因此,它是一种应用遗传知识和技术的医学实践。

作为面向社会的医学实践,遗传服务只是在20世纪下半叶人类与医学遗传学的理论和技术发展到一定阶段,与医学实践结合,达到了可以认真干预人类的生、老、病、死之后才开始出现的。在干预的过程中,人们发现有许多伦理、道德和社会方面的问题需要研究和解决,从而逐渐形成一门新的伦理学分支,即遗传伦理学(genetic ethics)。遗传伦理学是规范遗传服务中人与人之间的关系和行为的学科。

生、老、病、死一直是哲学家、宗教家和医学家自古以来就关注和试图解决的大问题。但哲学家大多从人生哲学,即人生的意义和对人生的态度方面思考问题;宗教家则进而考虑怎样摆脱生、老、病、死带来的烦恼和痛苦。道家主张随遇而安,佛家试图从根本上灭苦。但宗教信仰者追求的大都是精神方面的自我解脱,很少积极干预他人的生、老、病、死过程。真正面对和积极干预人类生、老、病、死的是医学家。他们在自己长期的实践中逐渐形成了一套道德伦理规范,即医学伦理学。

至于把遗传学知识用于人类自身的想法和实例,在古代东方或西方都可以找到。例如在我国,《后汉书》上曾记载,冯勤的祖父因身材矮小而"自耻短陋",于是给儿子选娶了一个高大的儿媳,结果生出了"长八尺三寸"的冯勤的故事。这是我国古代把遗传知识和手段(选型婚配)用于促进性优生,即个体性状的优化的实例。不过总的说来,这只是极个别的例子,因为当时遗传的概念还非常原始,干预的手段也极为有限。

相比之下,人类对自然界的遗传干预要积极得多。自古以来人类便不断自觉或不自觉地应用遗传学知识来改良植物、动物的性状,使之为人类利用和服务。育种和驯化已经成为人类文明史的一部分。而且在处理人和自然的关系上,我国古代很早就已经萌发了生命伦理学(bioethics)的思想,如提倡珍惜动物的生命。"君子之于禽兽也,见其生不忍见其死,闻其声不忍食其肉。"的名言就是我国古代这种思想的代表。至于医学伦理学,包括有关行医之道和医患关系的论述,中外历史记载俱多。但遗传伦理学只是在20世纪遗传服务普遍开展之后才逐渐形成。

真正地把遗传知识用于人类自身,即开展遗传服务始于20世纪50年代,当医学和人类遗传学有了长足发展之后。首先是人类细胞遗传学有了突破性进展。在这以前,人们一直不知道自己的染色体的正确数目,也不知道染色体的数目和结构异常可以引起疾病。20世纪50年代以后,由于组织培养、秋水仙碱、低渗处理等技术的发展和应用,瑞典的 A. Levan 和华裔学者 Joe Hin Tjio(蒋有兴)才报告了人类体细胞染色体的正确数目是46。此后,一系列的技术进展,如把外周血淋巴细胞作为制备染色体标本的培养对象、植物凝血素(PHA)作为淋巴细胞转化刺激剂的应用以及空气干燥制片法的发展,都大大促进了临床细胞遗传学检查的普及和许多染色体综合征的发现。20世纪70年代初显带技术的发展更使人们得以根据带型鉴别每一条染色体。它不仅解决了染色体的识别问题,还能在染色体上区别许多区和带,为深入研究染色体的结构异常及其定位创造了条件。此后发展起来的高分辨显带能在伸展得更长的染色体上展示更丰富的带纹,有助于发现更细微的染色体异常,并使其定位更加准确。再加上荧光原位杂交等技术的应用,染色体疾病的诊断服务更趋完善。

遗传服务成为可能的另一重要原因是分子遗传学技术的迅猛发展。20世纪70年代以后,重组DNA技术、DNA体外扩增技术和测序技术使人们认识的致病基因愈来愈多,可用于家系连锁分析的基因标记愈来愈丰富,这就为在家系中分析其成员是否带有致病基因,即进行基因诊断创造了条件。自 P. Harper 1983 年首次应用连锁分析于 Duchenne 型肌营养不良(Duchenne muscular dystrophy, DMD)基因携带者的检出以来,能通过连锁分析和基因突变的直接检测进行诊断与产前诊断的遗传病不断增加,基因检查现已成为遗传服务的主要内容之一。

与此同时,基因克隆、细胞培养、显微操作等技术的发展使人们想到,可以用引入正常的基因的方法来取代突变基因,达到治疗遗传病或肿瘤的目的。从20世纪90年代开始,基因治疗逐渐发展成为

一种临床遗传服务，尽管规模还比较有限。

从遗传服务的发展历程可以看出，遗传学技术的重大突破和应用对遗传服务的普及和内容的充实起着关键性的作用。可以预见，随着人类全基因组测序的完成，基因组学、生物信息学和各种高通量检查新技术的发展与普及必将大大拓展遗传服务的内容。例如，随着新一代测序技术的出现，如果个人全基因组测序的费用降到5000元甚至更低，可以设想，在不远的将来，当人们走进医院就诊时，或将首先出示一张载有个人全部遗传信息或基因检查结果的芯片。它将为医师提供就诊者整个基因组的情况，特别是与疾病和药物敏感性相关的基因的情况，从而为医师了解病因、病情与选用药物提供有用的参考和依据。

二、现阶段遗传服务的内容

在前面已提到，遗传服务的内容将随遗传学知识与技术的进步而不断扩展。在现阶段，医学遗传服务主要从事遗传病和与遗传有关的先天畸形的诊断、治疗和预防，服务的内容包括如下几个方面：

（一）遗传咨询

遗传咨询（genetic counseling）又称遗传商谈，指医师通过询问和检查，收集有关的家庭成员患病情况并加以分析，然后回答患者或询问者提出的各种问题。这些问题主要是：疾病是否是遗传病？如果是，会不会传给下一代？已有一个患儿时，再生出患儿的风险有多大？疾病的预后如何，能否治疗或预防？这些询问者关切的问题不仅与疾病的诊断治疗有关，还常常影响到咨询者本人或其亲属作出结婚和生育方面的选择。

为了正确地回答咨询者提出的问题，咨询医师首先需对疾病作出明确的诊断。为此常常要进行一系列的临床化验和遗传检查，包括细胞遗传和分子遗传学检查。这种检查常常不限于询问者一人，还涉及有关亲属。其次，咨询医师还需详尽地收集患者家族成员的情况，绘成系谱图，以确定疾病是否遗传及其遗传方式，并与检查结果相互印证。在这之后，咨询医师即可根据遗传方式、家系资料和检查结果，对疾病在亲属中的再发风险率进行估算。最后，医师应对该病的预后、防治的可能性提出意见，供询问者作出决定时参考。

从以上的介绍可知，疾病的确诊，遗传方式的确定和风险的计算是遗传咨询的三个最重要的环节，而每一个环节都要求提供家系成员的情况，有时甚至提供检查用的标本，因此都可能引发一系列的伦理学问题。

（二）遗传检查

遗传检查（genetic testing）是指对患者的遗传物质进行诊断性检查。这主要是为临床诊断和遗传咨询服务。按检查的内容，遗传检查可分为细胞遗传学检查和DNA检查。前者主要检查染色体，重点是诊断染色体疾病和肿瘤的染色体异常，以及辐射所引起的染色体畸变；后者主要检查相关基因，即通常所说的基因检查，用于基因病和肿瘤的基因异常诊断。DNA检查和细胞遗传学检查，尤其是前者，还可以用于法医学上的个人识别与亲子鉴定。

遗传检查还广泛应用于各种生物医学研究，包括遗传流行病学调查、药物遗传学研究等。

（三）基因治疗

将外源性基因引入细胞，以改变细胞或个体的表型来达到治疗的目的，称为基因治疗。通常是用正常基因取代突变基因来治疗遗传病、肿瘤或其他疾病。如将正常基因引入二倍体体细胞以纠正基因的缺陷，称为体细胞基因治疗（somatic cell gene therapy），而将正常基因引入生殖细胞、受精卵或早期卵裂细胞称为生殖细胞基因治疗（germ cell gene therapy）。后一种疗法引入的基因将一代一代地传递下去。基因疗法，尤其生殖细胞基因疗法在应用目的与后果方面有一系列的伦理和社会问题值得关注。

（四）辅助生殖

由于多种原因，大约10% ~ 15%的育龄夫妇是不育的。但现今已经可以采取一系列的辅助措施，以达到怀孕和生育的目的。目前较为普及和成熟的技术包括人工授精、试管婴儿（即人工体外受精和胚胎移植）。而在未来还可能采用体细胞核移植，即所谓人的克隆技术。

与自然生殖不同，在辅助生殖（assisted reproduction）过程中需要人为的选择或决策。例如在移植胚胎时，需决定用哪一个胚胎，什么时候进行移植等。其次，由于生殖细胞，即精子和卵子的供体可能不是夫妇双方本人，而对试管婴儿而言，还存在代孕或代孕母亲（surrogate mother）的问题，这样就会产生一系列家庭伦理问题。再次，辅助生殖过程提供了多种选择的机会，如选择供精者、供卵者、选择体细胞核的供体以及代孕母亲等。这些选择从本质上说是一种遗传选择，本应只以遗传健康为标准，但实际可能不是这样，因而同样可能引发一系列社会伦理学问题。

在整个卫生保健体系中，如果按提供服务的时段划分，还可区分为婚前、孕前、产前、新生儿期和儿童期的遗传服务。如婚前、孕前和产前遗传咨询，产前或新生儿期遗传筛查等。

（五）广义的遗传服务

上面介绍的几个方面的遗传服务限于解决个体和后代的健康与生育问题，或生、老、病、死问题。它仅仅是一种医学服务，我们可以把它称为狭义的遗传服务，或简称为医学遗传服务。在未来，用遗传学手段延长人的寿命也可归入医学遗传服务之列。

除此以外，遗传服务还可以扩大到社会生活的其他许多方面，我们把它称为广义的遗传服务。例如，在健康和人寿保险、就业、就学，户籍管理和个人识别，保密、防盗和社会安全等许多领域，都可以使用遗传技术进行鉴定和鉴别，并根据其结果作出选择与决策。例如，保险公司无疑希望了解被保人的遗传健康资料，如带有什么致病基因或易感基因、预期寿命几何等，以便对其健康和寿命作出预测，从而确定保险金额。同样，雇主也可能希望对应聘人员和员工进行身体、智力和性格方面的遗传预测，以便作出聘用方面的选择。同理，学校或老师、博士或硕士生导师也可能希望了解学生的身体、智能和性格等与遗传有关的资料，作为招生或培养的参考。政法部门可能希望建立个人的遗传资料数据库，以便在必要时用于个人识别和罪犯追踪。然而人们有理由担心，这类应用不当时可能导致基因歧视或种族歧视，个人隐私的丧失，同时也担心现阶段这种预测的可靠性，以及被简单化和滥用的危险。

为此美国国会在1996年通过了“健康保险流通及责任法案”（Health Insurance Portability and Accountability Act）禁止保险业者把预测性检测结果阳性，但尚无临床症状的人排除在保险之外。又于1999年通过“健康保险与就业反遗传歧视法案”（Genetic Non-Discrimination in Health Insurance and Employment Act），并于2000年由总统签署一项行政命令，禁止在任用和晋升雇员时使用遗传信息，以支持政府保护遗传信息、禁止受雇歧视的政策。

最后，还有一些违背常理的遗传服务领域，即把遗传学知识与技术用于发展基因武器和进行基因战争。理论上，通过遗传重组技术发展高毒性的或使神经瘫痪的病原体或毒素，或根据不同种族的遗传易感性发展有针对性的生物武器是可能的。但在这些方面，估计有众多的人反对提供这种服务。广义的遗传服务及其引发的社会伦理问题将不在本章讨论之列。本章讨论的只限于医学中的遗传服务。

三、遗传服务的特点

遗传服务虽是一种医学服务，但与一般医学服务仍有很大的不同。这是因为①个体的遗传物质即其基因组终生不变，遗传物质异常引起的疾病如遗传病也具有终生性；②在现阶段，许多遗传病还缺乏有效的治疗。药物、手术等一般不能改变或矫正遗传物质的异常。例如染色体病迄今仍没有有

效的治疗，而就大多数单基因病而言，治疗也多限于改善临床状况，根治还是未来的事；③由于异常发生在遗传物质，并可传递给下一代，这就使遗传咨询、遗传病的检查和处理不仅关系到患者本人，也牵涉到其他家庭成员和亲属；④上述咨询和检查获得的遗传信息不仅涉及患者个人，对家庭其他成员，包括下一代的健康也具有预测意义，因而可能引发一系列的心理的和伦理的问题。正是遗传性疾病的终身性、难治性和可遗传性等特点，要求医务人员在提供遗传服务时，充分考虑到服务对个人、家庭成员甚至社会可能产生的影响。

四、遗传服务的目的

在现阶段，遗传服务的首要目的应当是为遗传病和其他疾病的临床诊断、治疗和预防服务，同时它客观上也为防止遗传病患儿的出生提供了依据和手段。诊断和防治的对象主要是与染色体或基因异常有关的疾病。

作为以人为对象的医学服务，遗传服务与遗传学知识与技术用于选育或改造动植物品种不同。在充分考虑和认识长远后果，并达成社会共识以前，遗传服务不应用来改进人类正常基因或基因组的结构；不应用来改变一个民族、国家或群体的基因池的构成；也不应在尚未充分论证其有效性以前，开展任何有关个体智力发育、健康状况、对疾病，尤其是多基因疾病和肿瘤易感性的预测性遗传检查，以达到盈利的目的。

五、遗传服务遵循医学伦理学的一般原则

作为一种医学服务，遗传服务应遵循医学伦理学的一般原则。这些原则的核心是尊重（respect），即尊重个人的自主权（autonomy）、知情同意权（informed consensus）以及隐私（privacy）和保密（confidentiality）等。就遗传服务而言，是否要求和接受何种遗传服务，其后作出何种生育或人工流产的决定，都应由当事人自主决定；而遗传咨询和检查产生的信息和结果是否公开，也应由当事人自主决定。其他应遵循的医学伦学理学的原则还有：有利或行善原则（beneficence）、无害或不伤害原则（nonmaleficence）、公正（justice）或公平对待原则也应在遗传服务中贯彻。

鉴于医学伦理学在医学院校已作为一门独立学科讲授，并有相应的教材，读者可参考有关的书籍和文献资料。但如上所述，遗传服务在对象、方法和后果方面有一系列的特点。因此在遵循医学伦理学一般原则的同时，还应考虑遗传服务的特殊性，尤其是遗传病及其检测不仅涉及个人，还涉及家庭成员，而检测的结果和获得的信息与后代和亲属也可能有重要意义，因而可能引发更多的伦理问题。

六、发展中的遗传伦理学

在讨论和贯彻遗传伦理学的原则时有两点值得注意，这就是伦理观念的多样性和伦理观念总是处于发展过程中。

伦理学观念的多样性　人们的伦理道德观念是由他们所处的历史环境，包括社会经济条件、历史文化背景和宗教信仰决定的，因而也是随着社会的历史发展而发展变化的。因此，不同民族、社会在伦理道德观念的差异是自然的，甚至是必然的。就遗传伦理学而言，当前各国在对遗传服务、计划生育与优生、产前诊断、胚胎和克隆技术的应用等一系列问题上的认识、态度和政策就有很大的不同。此外，即使是同一社会，由于宗教信仰等不同，人们在对待遗传服务的态度上也可能不同。因此，要就诸如产前诊断的适用范围、产前诊断后人工流产的指征，辅助生殖措施应用过程中是否应该和如何进行供体选择等达成共识还有待时日。

遗传伦理学是发展的　既然人们的伦理道德观念是随着社会经济文化发展变化的，遗传服务的技术和内容也是发展的，那么遗传伦理学的内涵研究重点和观念也在变化中。每一次技术方面的突

破，尤其是遗传学技术的重大突破，都可能产生新的遗传服务内容，同时也引发新的伦理道德问题或加剧原有伦理道德问题的困惑。20世纪后半期细胞遗传和分子遗传学技术的突破及随之兴起的遗传检查是这样，而基因治疗和辅助生殖技术的迅速发展也是这样。

人类基因组计划与遗传伦理学　本世纪初人类基因组计划的完成和大规模高通量测序技术的进步，不仅带动了系统生物学和生物信息学迅猛发展，也使与医学有关的遗传检查和基因诊断日益普及。当前遗传检查正从染色体疾病和单基因遗传病为主的诊断检查过渡到心血管疾病、肿瘤等常见复杂疾病相关基因的检查。这一重要发展固然是临床遗传学的福音，但同时也使遗传选择和基因歧视问题更加尖锐；而大量个人遗传信息的获得也加剧了保护隐私权、知情权的困难。可以设想，当胎儿全基因组测序变得非常容易和廉价时，父母和社会必将面临更加艰难的选择与伦理困惑，因为在每个胎儿通过检查都可能发现一些（例如8～10个）致病的，“不良的”基因，而遗传工程技术的发展又可能诱惑人们追求完美无缺的婴儿。可见，随着新技术的不断出现，现有的伦理问题尚未解决，新的问题又将出现。遗传学家与医务人员对此应有充分思想准备，及早开展对新技术引发的伦理学问题的探讨。

医学遗传学领域的伦理问题甚多，诸如获得遗传医学服务的社会公正性、遗传咨询、遗传选择的自主决定权和选择权、个人遗传信息的保密和是否与家庭风险成员分享、遗传咨询建议的指令性和非指令性、产前诊断的范围和应用，包括性别的诊断和选择以及在基因治疗和医学遗传研究中的伦理问题等。但本章将以医务人员和医学生为主要对象，分节介绍各项遗传服务，特别是与临床实践直接有关的各项遗传服务中值得关注的伦理问题。

第二节　遗传咨询中的伦理问题

一、体察咨询者的心态

伦理学问题实质上是人与人之间的关系和应遵循的原则问题。如，医学伦理学主要是研究医师和患者之间、医师和医师之间在医学服务中的关系和应遵循的原则问题。

前面已讲到，遗传咨询是医师和咨询者的遗传商谈。但这种商谈不是一般的商谈，也不完全像普通的门诊，而寻求咨询者与一般患者也有很大的不同。为此，咨询医师除了应遵循一般医患之间的伦理道德原则外，还应特别关注寻求咨询者的心态。咨询者通常包括如下几种人：①疑为遗传病患者，要求确诊；②已生下了一个患儿的夫妇，关心是否会再次生出患儿；③婚前男女，其中一方或其亲属为遗传病患者，他们关心婚后的子女是否会患病；而在一些迟发的遗传病，如 Huntington 舞蹈病，Becker 肌营养不良等，还关心自身是否会发病；④亲属中有遗传病患者，他们关心自身及子女患病的可能性。此外，询问者经常提出的问题还包括近亲可否结婚生育，肿瘤、精神病或其他常见病会不会遗传等。

由于遗传病的难治性和可遗传性，许多咨询者前来咨询时心存顾虑。这种心态源于：①一种羞耻感。正像《遗传咨询》一书作者富尔曼等指出的那样，不少家庭对出现遗传性疾病就好像出了什么丑事，甚至像犯罪一样，想方设法隐瞒。配偶双方有时甚至为此相互指责；②一种负罪感，尤其是生育了遗传病或先天畸形患儿的父母，他们认为是自己把疾病传给了子女，给他们带来了不幸。此外，还有一种对患病情况被宣扬出去的恐惧。

因此，咨询医师应该体察这种心情，并设法减轻咨询者的羞耻感、负罪感和恐惧。通常的做法是：①强调遗传病是疾病，不是性传播疾病，不必有羞耻感；②强调遗传病是这样的疾病，不论患者本人或他们的父母都没有任何过错；即使致病基因确由父母传递，但那是不以他们的意志为转移的事；③它也不是一种惩罚（更不是什么因果报应！），而父母也无须与自己的任何行为或过失联系在一起，自责或互相指责，或有任何道德或伦理方面的思想负担。

二、遗传咨询时应遵循的原则

遗传咨询作为一种医学服务，除应体察咨询者的心态外，整个过程都应根据遗传咨询的特点贯彻医学伦理学的基本原则。如：

（一）尊重隐私权

遗传咨询不宜在有无关人员在场的环境中进行。为个人的隐私权得到充分尊重，必要时咨询医师可以与前来咨询的夫妇、亲子分别谈话。这是因为遗传病不像感染性或其他疾病，只涉及患者本人，而家系调查不可避免要涉及亲属，如父母、兄弟、姐妹。除了信任医师以外，咨询者可能不愿其他人，甚至自己的配偶知道自己和家人的情况。笔者曾遇见一位母亲为了自己成年儿子的生育问题，要求医师代为详细了解儿媳一家的患病情况。这说明即使在家庭内成员或亲属之间，对有关遗传情况的交流仍然是敏感的。因此，咨询医师除尊重咨询人的隐私权，同时也应尊重有关亲属或成员的隐私权，为通过咨询获得的资料保守秘密，避免为他人、单位、雇主和保险商等利用，以利于家庭的和谐与稳定。

（二）自主和知情同意

遗传咨询本身应是自愿的。因此当咨询要求患者及其家系成员进行遗传学检查及临床化验时，也应贯彻自愿的，即知情同意（informed consent）的原则，让患者及有关人员充分了解检查的目的与必要性，并在合乎对患者或受检人有益无害的情况下，争取他们的主动配合。

（三）自主决定和非指令性的（non-directive）原则

遗传咨询和检查的结果有可能证实遗传病的存在，或计算出疾病的再发风险。如已证实或高度怀疑胎儿为21三体综合征患者，此时咨询医师应当向父母详细介绍疾病的原因、后果和预后，再生育时不同核型21三体综合征的再发风险以及胎儿产前诊断的风险率，和各种可能的处理办法等。但咨询医师不应代替父母作出任何处理的选择或决定，包括是否继续怀孕，是否作胎儿产前诊断，或是否人工流产等。咨询医师即使提出建议，都应是非指令性的，决定权最终应属于父母。至于医师应否建议和表达明显的倾向性，国情、人情不同，目前尚无共识，值得继续讨论。

第三节　遗传检查中的伦理问题

遗传咨询常常伴有一系列的检查，包括基因检查和染色体检查。目前我国广大临床医师对于能提供任何诊断或治疗信息的检查一般都持积极的态度，认为为了疾病的确诊，检查是必需的，但较少想到会引起什么伦理方面的问题。一个原因是，能开展的检查项目有限，获得检查服务的机会不多，当前多限于一些大的医疗机构才能提供。尽管这种状况在全国范围内尚无根本改变，但随着临床遗传学的需求和疾病基因组学技术的发展，致病基因的克隆和诊断将加速和日益普及。在一些大型的教学医院，连锁诊断或致病突变的直接分析已逐渐成为常规服务。在这种情况下，检查引发的伦理问题值得重视。

遗传检查按进行的时间顺序可分为胎儿产前检查，包括胚胎着床前检查、患者的诊断性检查和预测性检查，而按检查的目的可分为诊断性检查和科研性检查。

医师在进行遗传检查时应遵循尊重隐私、知情同意和对咨询者及相关成员有益无害的原则。然而在实践中，贯彻这些原则并不容易。就知情同意而言，考虑到我国广大患者的文化水平和科学知识，这需要咨询医师极大的耐心和具有用深人浅出的语言释解复杂问题的能力。就有益无害的原则而言，需要咨询医师能抵抗增加收入、发表科研论文等各种名与利的诱惑，摒弃各种于患者治病防病本身非必需的检查。在当前，我国的遗传检查，无论是染色体检查或基因检查都较昂贵，它可能成为“创收”的来源；而克隆新基因和发表高档次的研究论文又可名利双收，因而在实践中伦理道德的考

虑可能让位于利与名方面的考虑。

唯一可以证明这样做是正当的理由是:对患者全面深入的遗传检查和临床数据积累可能加深对遗传病的认识,在未来有益于患者的诊治。

遗传检查除应遵循上述原则外,还有一些特别值得注意的问题。它们是:迟发遗传病的基因检查;儿童的遗传病检查;胚胎的遗传检查;家庭风险成员的检查和一般个体的预测性基因检查。

一、迟发遗传病的检查

分子遗传学技术提供了在任何年龄识别突变基因的方法。乍一看来,这似乎有利于疾病的预防。然而对于某些遗传病,尤其诸如 Huntington 舞蹈病(Huntington chorea)和成人多囊肾病(adult polycystic kidney disease,APKD)这样的迟发遗传病,通过检查让患者在发病前就知道他是致病基因的携带者,究竟是利大于弊呢,还是弊大于利呢?

这些迟发性遗传病目前尚无有效的根治手段,甚至也无防止临床发病的有效措施。如果把阳性结果告诉检查者,那么,首先可能增加检查者的思想负担,使其一生的生活处于阴影之中。其次,迟发性遗传病的基因携带者常常在发病前已经结婚生育,并有可能已经把致病基因传递给了下一代。如果是这样,那么阳性检查结果可能引起携带者的负罪感和其他家庭问题。再者,如果检查结果被人知道,还有可能导致携带者在婚育、就学、就业、医疗保险等方面受到歧视。

因此,在进行迟发遗传病的发病前基因检查时,首先应进行遗传咨询,让受检人通过咨询充分了解疾病是否遗传、能否治疗以及传递的风险等,使之有充分思想准备接受检查的结果。经过咨询后,检查者可能会改变初衷,放弃基因检查。其次,任何检查都应征得受检人的同意,咨询医师不应为了谋利或其他的目的进行指令性的检查。最后,任何检查的结果都应当为受检人保守秘密。

二、儿童的遗传病检查

在遗传病家系中,如果风险成员是儿童,他们应否检查是一个十分困难的问题。这是因为,儿童无法实现知情同意的权利。而当检查的是无法治疗的疾病时,这种检查常常不仅对他们无益,还可能使他们遭到歧视。

站在胎儿或儿童的立场,迟发遗传病,尤其无法治疗的遗传病的预测只能带来负面的影响。这种检查为人工流产提供了依据,剥夺了胎儿出生的权利。而对已出生的儿童,这种预测结果可能导致家庭和社会的歧视,包括放弃治疗、不再抚养,尤其在经济困难的家庭和发展中国家,同时也剥夺了儿童将来自主决定的权利。

另一方面,父母可能希望对风险儿童进行预测,以便及早作出安排。在我国,这还涉及是否获准生第二胎的问题。而第二胎的出生可能使患儿更加受到忽视或歧视。咨询医师在这里处于一种两难的境地。是优先考虑胎儿和儿童的权利呢?还是优先考虑家长和社会的利益呢?困难在于,胎儿和儿童自身还无法作出检查与否的决定,即不可能行使自主决定权。由此可见,在涉及遗传服务,尤其是遗传检查和随之而来的遗传选择时,个体、家庭和社会的利益并不总是协调一致的。而父母、社会或在哪种程度上有权为胎儿或儿童作出决定,仍是一个可争议的问题。因此,一些遗传咨询医师建议,把无法治疗的遗传病的预测检查留待风险儿童成年之后,由本人自行决定是否进行检查。

在一些国家,由于传统或宗教的原因,可能更加重视个体的权利,包括自决权,因而对遗传检查和人工流产持保守的态度;而在另一些国家,特别是发展中国家,由于人口的压力和许多家庭无力照看一个遗传病患儿,可能对遗传检查与人工流产持更积极的态度。

三、胚胎的遗传检查

(一) 着床前的遗传检查

由于试管婴儿的日益普及和遗传检查技术,特别是着床前诊断技术的提高,在将胚胎植入子宫前即可采取早期胚胎的单个或少量细胞进行染色体和基因检查,从而避免植入有染色体异常或致病基因的胚胎。如,一些国家已批准对有家族性腺瘤样息肉病(FPA)、遗传性乳腺癌、前列腺癌和老年性痴呆等风险家系的胚胎移植前检查。可以预见,此类检查项目还会不断增加,然而一些生殖伦理学有关人士则认为,为了保证不携带任何一个致病基因而毁弃大量胚胎有悖伦理。这种筛查如果不严格限用于有严重遗传性疾病或肿瘤夫妇的着床前胚胎,则将导致严重的伦理困惑并可能陷入按父母意愿“定做婴儿”的误区。

(二) 胚胎的遗传检查

如果上述遗传检查不限于着床前发育早期的胚胎,那么涉及的伦理问题就更加复杂和严重。因为检查的目的仍是避免患儿出生,这除了涉及入选检查基因的标准外,还有一旦检查出疾病相关基因时应如何处理胚胎的问题。如①携带什么疾病相关基因者可以人工流产,②多大胚龄的胚胎可以人工流产、弃置不用或销毁?③谁有这方面的决定权?

在这里,一个关键的、也是难以达成共识的生命科学和伦理科学的问题是,作为一个人,他的生命从何时开始?是从神经系统开始发育算起,还是神经系统发育到某一特定阶段后算起?抑或应从胚胎有感知或自我感受算起?如何得知胚胎是否有感知和感受?这些问题比医学家曾经历过的、有关死亡判定标准的辩论要复杂得多。因为判断心跳停止、呼吸停止或脑死亡要客观和科学得多。

此外,即使人的生命从何时开始的问题,能够较科学地解决或达成共识,出于政治经济、社会文化、宗教信仰等原因,人们对待胚胎的遗传检查的政策和态度也还可能不同,或不断发展改变。英、美等发达国家对待人类胚胎、胚胎细胞用于诊断、治疗和科研的政策的多次调整就是明显的例子。

四、家庭风险成员的检查

在遗传病家系中,除患者外,其他风险成员也可能自己要求检查,以了解自身是否是致病基因携带者。但通常更可能的是:在进行遗传咨询,尤其是家系调查时,要求未患病同胞和亲属提供标本进行染色体或基因分析。这确实可能引起一系列的伦理学问题,其中包括咨询者的心理负担、对亲属的亏欠感或负罪感以及引起未患病同胞或其他风险成员对检查结果的恐慌,因为他们有可能被诊断为致病基因的携带者。因此,前面提到的自主决定、知情同意、保密和有益无害等原则,也完全适用于风险成员的遗传检查。

在迟发性遗传病发现风险成员携有致病基因时,应否把检查结果告诉本人也是个问题。因为他本人可能并未要求检查,或者对接受检查结果未做好思想准备。他有知情权,也有不知情权。

风险家庭成员的遗传检查有可能产生上面一些负面影响。然而,从家庭和社会预防遗传病患儿出生出发,应该承认这样的检查又是有益的,甚至是必需的。例如,一个 Huntington 舞蹈病患者的成年子女如检查结果证实为致病基因携带者,他可以选择不结婚,婚后不育,或采取产前诊断、异源人工授精等措施,以避免带有致病基因的胎儿出生,而这对家庭和社会都有好处。

五、一般个体的预测性基因检查

近些年来,由于基因组学及有关技术,特别是高通量 DNA 测序、基因芯片技术和生物信息学的发展,人们识别和分离克隆的基因愈来愈多,对它们在人体各种生命过程、代谢通路和疾病中的作用的

研究也日益增多。这些研究除发现或证实了不少单基因性状和疾病外，还提示另一些基因可能与一些复杂疾病或性状有关。验证和阐明这种相关性的研究称为关联分析（association analysis）或关联研究，它在当今遗传医学，尤其是对多基因疾病、先天畸形和肿瘤的遗传研究中占了很大的比重，也获得了一些阳性结果。

尽管这些阳性结果绝大多数还是非常初步的，仅仅表明某个基因与某种多基因疾病或性状有关，但已有人把这些只是可能有关的基因用于一般人群的复杂疾病、健康状况与寿命、生殖与发育、甚至智能与天赋的预测性检查。在我国，这种针对既非患者、也非有某种遗传风险个体的，有关其生、老、病、死的预测性基因检测的兴起，除上述基因组科学的进步提供了技术条件外，还因为：①改革开放使一部分人富了起来，使昂贵的预测性基因检查有了市场；②媒体的报道和科普宣传在增加群众遗传学知识的同时，常使用一些通俗易懂、但也容易误导的词汇，诸如长寿基因、智力基因、肥胖基因、糖尿病基因、高血压基因等，使人们误认为极为复杂的生命现象、多基因疾病或性状也可以通过上述基因的检查来预测；③生物医学公司或医疗机构的营利动机和④科研人员面临的研究成果转化压力，也起到了重要的或推波助澜的作用。然而，抛开支持和提供这类服务者的动机不论，这类基因检查的问题是科学性不足和与之有关的伦理方面的欠缺。

科学性不足主要是：少数基因与复杂疾病的关联远远不足构成诊断或预测的基础。关联只能说明两者有关，既未说明关联是直接的还是间接的（直接影响基因功能，还是仅仅与其他功能基因处于连锁不平衡状态），也未说明其机理，更不能说明多个有正负效应的相关基因同时作用时的效应。当人们就一种代谢途径或一个通道上的几个基因尚不能阐明其相互作用，提出有根据的互作模型、适当的算法和软件加以评估验证，又怎能就涉及众多基因的疾病或性状作出科学的预测？这种科学性不足是因为现阶段对各种复杂疾病或性状的遗传基础既不清楚，也无法估计表观遗传和环境因素的影响，仅凭少数基因关联还远不足以作出预测的原因。

当前遗传咨询中对多基因疾病或畸形再发风险率的预测，目前也还十分粗略。即便是针对有遗传负荷的风险个体，预测也是基于对患者家系资料的经验风险率的估算，而非基于基因检测的结果。

以任何一种常见的多基因疾病为例，无论是糖尿病、还是唇腭裂、或是精神分裂症，究竟有多少基因参与其发病？它们各自起的作用大小如何？有无主基因？作用的机理如何？这些基因间的相互关系或作用如何？除基因外，还有什么因素影响基因的表达？对于这些问题目前知之不多，无法回答。而就精神分裂症这样重要的常见病而言，至今甚至还没有一个可供诊断参考使用的基因。在这样的情况下作出的预测可信吗？因此，现阶段在没有阳性家族史的情况下，用基因检测来预测入学儿童或入伍新兵是否将会罹患精神分裂症是不可能的，更何况预测其寿命、智能了。

科学性不足本身即意味着不符合伦理原则。①首先是它不符合有益无害的伦理原则。因为预测本身科学性差，准确性更是可疑，提供者尚难自圆其说，难以对受检者有益；②不确定的检查结果或“阳性结果还可能引起受检者的忧虑和恐慌，故有害无益；③目前这类检查大多收费高昂，是一种无益的经济负担；④从社会卫生资源的分配利用而言，它也不符合公平公正的伦理原则：一方面对许多迫切需要诊治的遗传病、白血病和肿瘤等患者因医保投入不足而不能获得应有的诊治，另一方面又将大量人力资金用于科学论据不足的昂贵检测，这就难怪许多有识之士均不提倡、不赞成此类基因检查。

对待一般个体和群体的基因预测性检查的正确的做法是：①任何临床基因检查项目都应对其科学性、实用性、可验证性和局限性进行充分论证；②严格审批此类检查项目进入市场服务，尤其是由非医疗机构以盈利为目的的服务；③提高媒体在涉及遗传医学、人类基因组学进展方面的报道和科普文章的科学性；④加强基因功能、表达调控、基因互作、基因与表型性状、基因与疾病，尤其是多基因性状和疾病的研究，从而为未来科学的无遗传负荷个体的预测性基因检查奠定坚实基础。

第四节 基因治疗中的伦理问题

将正常基因引入二倍体体细胞以纠正致病基因引起的缺陷,称为体细胞基因治疗。它比较简易可行,也不引起严重的伦理学问题,因为引入基因不致影响到下一代。不过"治愈"的患者仍是致病基因的携带者。

生殖细胞基因治疗则不同。由于治疗基因引入了生殖细胞或受精卵,它可以一代一代地传下去,从而改变子孙后代的遗传结构。这就提出了一系列的社会伦理问题。其中主要是:现阶段基因疗法能否保证安全和有效性,什么是生殖细胞基因疗法的适用范围或指征,而伦理问题的核心是人类应不应该改造自身的遗传结构,或者为了什么目的进行这种改造才是合理的。

从1990年美国 Blease 和 Culver 等对腺苷酸脱氨酶缺乏症进行基因治疗并获得成功后,在全世界曾掀起了基因治疗的热潮。直至2000年5月,经各国政府批准的临床试验达到425项,其中大半涉及肿瘤(65.6%),其次为单基因遗传病(12.9%),其余为感染性疾病等,约30余名患者已接受过基因治疗。但是,由于新闻媒体的过分渲染和商业机构的过早卷入与炒作,在基础研究尚不充分的情况下,一些临床试验仓促上马,而且大多偏重于具有商业价值的肿瘤的基因治疗。科学家经过冷静地反思,1995年以后基因治疗开始降温,美国国立卫生研究院也采取了严格考核临床试验,暂停支持新项目等一系列措施,而强调回归到对疾病的基础研究方面。然而,从长远来看基因治疗仍有光明的前景。

一、基因治疗的安全性与有效性

在一些遗传病如 ADA 缺乏症、血友病 B、家族性高胆固醇血症,已经证明了体细胞基因治疗的有效性和安全性。但对大多数还处于临床实验治疗阶段的疾病,这二者仍有待充分证明。在这方面,对治疗方案将逐个地进行严格的考评。一般来说,对肿瘤基因治疗的批准要宽松一些。但实际上除少数遗传性肿瘤以外,多数肿瘤涉及的基因众多,发病机理比较复杂,在对基因的表达调控和相互作用还知之不多的情况下,基因治疗在相当一段时期还不能替代手术或放疗化疗而成为主要的治疗手段之一。

二、现阶段的生殖细胞基因治疗

(一) 生殖细胞治疗的优点

和体细胞基因治疗不同,生殖细胞基因疗法的优点是:①引入基因可以一代代地传下去,治疗的效果带有终身性和可遗传性;②一些基因只在胚胎发育的特定时期才表达或起作用,这些基因异常引起的遗传病只有在胚胎发育期才能防治;③一些神经系统的遗传病治疗可能由于血脑屏障等原因,只有采用生殖细胞疗法在屏障形成以前才会奏效。

(二) 生殖细胞治疗条件还欠成熟

在技术上生殖细胞疗法与转基因动物操作类似。而近年来在动植物转基因已取得很大成功,多种转基因动物相继问世,标志着人类也将进入了转基因时代。然而在现阶段,转基因技术用于人类生殖细胞,似为时过早。这既有技术上和认知上的原因,也有社会伦理学上的原因。

1. 技术上的原因　概括地说,现阶段人们对真核细胞基因的表达调控知之不多,进行调控的手段还不成熟。因此,如何使引入的基因能恰如其分地表达,还需有新的技术突破。换言之,治疗的有效性问题还需认真研究。

2. 认知上的原因　人们对基因之间相互作用和疾病之间的相互关系知道得仍然太少。例如,引入基因是否会激活或抑制其他基因,包括肿瘤相关基因,从而引起意想不到的后果,目前还不能肯定。其次是致病基因相互关系常常出人意料。如,地中海贫血有严重的后果,但奇怪的是引起地贫的基因

传播竟如此广泛，在某些地区甚至每5个人就有一个携带者。后来才知道，地贫的杂合子具有一种优势，它们具有抗疟疾的能力，地贫和疟疾地理分布的一致性证实了这一点。同样，镰形红细胞贫血症、G-6-PD缺乏症的基因携带者也具有上述优势。这些例子表明，在采取任何基因治疗措施时应十分慎重，尤其是采用生殖细胞基因治疗时。

3. 伦理和社会方面的原因　除了治病以外，人们对应否为了其他目的而引入外源基因，从而改变或“改良”自身和后代的基因组结构尚未取得一致意见。

支持者可以举出许多理由。如，既然在进化过程中人类的基因组是不断改变的，只不过这种改变是被动的、渐进的，那么人类为什么不能主动地、更快地改造自己的基因组结构，使人类变得更健康，更聪明或更漂亮呢？也就是说，为什么不能通过引入外源基因，达到正优生学的目的，达到基因组的优化呢？

反对者当然也有许多理由。他们人数众多，是对基因技术及其产品，包括转基因动植物和转基因食品等持审慎态度者。首先，一部分人认为，上帝创造或安排的一切都不宜轻易去改变。这在多数不相信上帝的科学工作者看来是可笑的。然而，如果把“上帝创造的”改为“长期进化形成的一切”不宜轻率去改变，则不无道理。因为正如前面所说，除了防病治病以外，①现阶段的技术还不够成熟，对基因组的认识还不充分，还不能保证对基因组的优化或“改良”成功而无副作用；②人们还没有公认的“优”与“劣”的标准，也就是对改造的标准，对未来理想人类的设想还无共识；③人类的遗传多样性既有生物学上的意义，同时也有社会生活的价值。一个简单的例子是，给小鼠引入生长激素的基因可以使一般小鼠发育成“超级小鼠”。在一定发育阶段，给以生长激素或进行体细胞基因疗法可以治疗侏儒症，也可以使普通人身高增长。但是如为了子女长得更高而向生殖细胞引入类似的基因就会有一系列问题，包括什么高度是人类的理想高度？理想标准高度是否会随历史发展而改变？引入基因的表达是否会引起巨人症或肢端肥大症？其他的考虑还有：一个群体由于具有多样性可能更能适应自然环境的变迁，而一个社会可能由于多样性而更富生活情趣等。

三、基因治疗的指征和条件

因此，在现阶段，人类基因治疗的应用似应限于：

1. 遗传病治疗，尤其是严重的、现阶段尚无有效治疗的遗传病，以及其他疾病，包括肿瘤和艾滋病等尚无有效治疗手段的严重疾病。

2. 治疗技术比较成熟，引入基因表达调控手段比较有效，且经动物实验证明治疗有效的疾病。

3. 基因引入不会激活有害基因如癌基因和抑制正常功能基因。

与此同时，应坚持现阶段基因治疗只针对疾病，而不应该用于：

1. 促进性优生的目的，包括未经充分界定的“优化”，“改良”或含义尚不清的“遗传素质的提高”。

2. 政治或军事的目的，即通过改造遗传结构而达到控制某一个体、群体、民族的目的，或用于发展基因战争等。

第五节　辅助生殖中的伦理问题

一、辅助生殖技术(assisted reproduction technology，ART)

过去不育只能通过领养来解决，但现在已可以采用各种辅助生殖技术来获得子女。这些技术包括人工体内受精(artificial insemination，AI)，体外受精(*in vitro* fertilization，IVF)和胚胎移植(即所

谓试管婴儿)，以及与之有关的代孕母亲。最后，还有尚有争议的无性生殖，即所谓克隆人的技术等。

各种辅助生育所引起的伦理、法律和社会问题大致可分为两类。第一类是由于采用了其他人的生殖细胞或生殖器官(如子宫)所引发的家庭伦理或法律问题。第二类是辅助生殖手段提供的遗传选择的机会所引发的选择依据、选择标准和选择后果的问题。

我国卫生部于2001年颁布了人类辅助生殖技术管理办法和人类精子库管理办法。这是我国开展辅助生殖的行政规范和依据。以下侧重介绍辅助生殖中的伦理问题。

二、人工授精的家庭伦理和社会问题

(一) 家庭伦理问题

人工授精在我国已十分普遍，它用来解决男性不育的问题。采用丈夫精子(artificial insemination by husband，AIH)不存在严重的伦理法律问题，最多只有性与生育分离的问题；但使用供体精子(artificial insemination by donor，AID)则有一系列的问题需要考虑和明确。有关精子库和人工授精的实施有明确的条例规定。而伦理方面目前已经比较明确的是：①供精者只是提供精子或遗传物质，不能成为孩子的父亲，以避免引起伦理和法律问题。②尽量维护受精者的家庭稳定和避免家庭伦理问题的发生，因而应为当事人即供精者和受精者保密并不应让孩子知道谁是供精者等，以利于孩子的健康成长。③不应过多地使用同一供体的精子。按我国规定，一名供精者最多只能提供精子给五名妇女受孕，一则避免产生众多的"同父异母"兄弟姐妹，而这种可能性确实出现过，因为2009年在美国曾发现一名供精者在其一生中不断地多次供精；二则避免同一供体精子的多次使用可能导致的群体中相同基因单倍型频率的增加，后者既不利于群体的多态性，又可能使隐性致病基因纯合化的机会增加。不过，人工授精与自然受精相比始终是少数，负面影响的大小还需进一步评估。

(二) 实施原则

出于上述伦理和家庭稳定方面的考虑，在实施人工授精中应当贯彻如下原则。①夫妇双方自愿，并提出申请；②院方严格控制指征，包括要求出具不孕症的证明等，并不应为未婚男女、寡妇、身体健康而又非不孕症患者实施人工体外受精；③供精者知情同意，有妻室者还应征得妻子同意；④供精者与受精者互盲，与后代互盲；⑤实施人工授精操作的医务人员与供精者互盲，与后代互盲。所有这些保密要求主要是保护受精者的利益，并有利于孩子的健康成长。当然受精者也应对后代保密。

三、体外受精和胚胎移植

通常所说的"试管婴儿"，指的是取出成熟的卵细胞后在体外受精，并在卵裂至4～8个细胞时将幼胚移植回子宫。它可以解决输卵管闭塞所引起的女性不育，也可以通过卵细胞胞质精子注射(intracytoplasmic sperm injection，ICSI)解决男性精子不能穿卵的问题。目前已有数以千计的"试管婴儿"出生并已长大成人。

体外受精的家庭伦理问题

试管婴儿的家庭伦理问题主要是因孩子有多个父母和利用代孕母亲引起的。所谓多个父母是指：提供精子或卵子的遗传父母、生后的养育父母、兼有两种身份的完全父母以及提供子宫的代孕母亲。

因此，试管婴儿可以①只与父亲有血缘关系；②只与母亲有血缘关系；③与父母均有血缘关系；④与父母双方均无血缘关系；⑤与母亲有孕产关系；⑥与母亲无孕产关系，即由代孕母亲孕产。

由于婴儿与父亲、母亲、代孕母亲在提供遗传物质、怀孕、生产和抚育三个方面可能出现的不一致

性，这就导致了比人工体内受精更加复杂的家庭伦理和法律问题。即以谁提供卵细胞、谁完成怀孕和生产、产后谁抚养就可区分遗传母亲，孕产母亲和抚养母亲等。

虽然传统观念认为有血缘关系，即认为遗传父母才是真正的父母，但从稳定家庭和有利人工授精技术的应用出发，大多数国家，包括我国在内都主张抚养教育的父母才是真正的父母，并从法律上加以确认。

从同样的考虑出发，大多数国家均主张对孩子保守遗传父母是谁的秘密，但也有少数国家（如澳大利亚、瑞典等）容许了解遗传父母的情况，从而有可能引起孩子与抚育父母之间关系的不稳定。

至于代孕母亲是否符合道德与法律目前仍然有争议。赞成者认为代人怀孕可以让不能怀孕妇女获得子女，应该受到欢迎；批评者则怀疑代孕者是出于商业动机而非怀着高尚的目的。我国政府规定医疗机构和医务人员不得实施任何形式的代孕技术。因此，除美国外，大多数国家都禁止代孕母亲。

在提供精子和卵子时也有人为的遗传选择的问题。如身心健康是对供精者、供卵者的基本要求。一些保证胎儿身心健康的特殊筛查选择也是合理的。如，不论精子供体是丈夫或其他人，都可检查其精子染色体是否存在有碍受精的微缺失等。此外还有代孕母亲的问题，因为代孕母亲提供了胚胎发育的场所或环境（子宫），她的身心健康无疑也十分重要。但不应对供精或供卵者，甚至代孕母亲提出超出生殖健康以外的要求（参见有关诺贝尔精子库一节）。

四、人工授精供精者的选择

除家庭伦理问题外，另一个比较复杂和引起争议的社会问题是供精者的选择问题。实际上，自人工授精技术开始采用以来就存在选择标准的问题，而且至今尚无一致意见。20 世纪中期开始的美国有关“诺贝尔精子库（Nobel sperm bank）”的争议和几十年后，即 20 世纪末国内有关“名人精子库”的辩论就是这种争论最突出的例子。

希望自己的孩子健康、聪明是作为父母的普遍愿望。但认为通过选择精子供体就能生出漂亮或智力超常的儿童则是一种奢望。它并不符合遗传学原理，也未客观评价个人成长中遗传与环境因素的作用。现今“诺贝尔精子库”或“名人精子库”的宣传，除少数例外，可能都出于商业动机而没有向公众讲清，现阶段什么是可能的，什么还做不到，以及这样的选择可能产生的负面影响，因而只起到了误导的作用。作为医务工作者，他们本应比一般公众更了解问题的实质。

（一）诺贝尔精子库不宜提倡

“诺贝尔精子库”或“名人精子库”之所以不宜提倡，其理由有遗传学的，也有哲学的或社会伦理学的。

从遗传学的观点，诺贝尔精子库不宜提倡的理由主要是：基因、染色体和生殖细胞的随机分离组合不能保证优良基因传递给子代。

1. 精子生成过程中由于基因和染色体的重新组合，不能保障“优良”基因传递给孩子，更不能像一些报刊所宣扬的那样“精子银行让你花钱订制下一代”。总的原因是，由一个初级精母细胞经历了两次减数分裂最终发育成四个精子的过程中，有多次基因和染色体的随机组合，这包括①发生在第一次减数分裂前期一对同源染色体的联会（synapsis）时染色体片段的交换（crossover）导致的基因的随机重新组合，这种交换平均多达 30 ~ 40 次，而且参与交换的染色体片段的位置、长度，即交换了哪些基因是随机的，不可预测和不可控制的；②联会时两条同源染色体形成的二价体（bivalent）在细胞分裂后期的分离是完全随机的。人类细胞有 23 对同源染色体，产生的染色体组合可高达 2^{23}（8 百余万）种；而第二次减数分裂时半二价体（half bivalent）两条染色单体的分离也是完全随机的。因此，在

减数分裂中发生的一系列随机重组事件好像多次洗牌,而洗牌后的分牌结果是难以预测的。因此,要保证特定的基因,如与决定高智力有关的基因传递给孩子是困难的。

2. 精卵结合的随机性也难以保证优良基因的传递。当精卵受精成为合子时,数以百万计的精子中哪一个与卵子结合,也是随机的和不可预测的。其结果是:形成的合子的基因型既不同于父母,也不同于其他任何人的基因型。因此,除同卵双生子外,每一个人的基因组成都是独特的,不重复的。这也是人类多态性的遗传学基础。

3. 每个人都有一些“不良”基因,诺贝尔奖获得者或“名人”也不例外。一个群体中的每一个体平均有8~10个突变基因。不排除诺贝尔奖获得者在某些方面有由基因或基因组合决定的天赋,如智力超群。这种超常的智力通常是多个基因决定的。然而,正同普通人一样,他们也有一些“不良”基因或突变基因,如疾病易感基因或致病基因,而且同样可以传递给下一代;其次,基因的“优良”与“不良”,正如前面已经提到,还需视具体情况而定,即放在特定环境和特定个体的基因组合中来判断。在现阶段,人们基本上还不清楚成千上万个基因之间的相互作用,也不知道什么样基因组合是最佳的组合。而即使是“优良”的基因,是否在各种组合与环境条件下都“优良”,同样是一个难以回答的问题。

以上是从遗传学的观点来看所谓“诺贝尔精子库”不宜提倡的理由。同样的理由也适用于对模特儿卵子的炒作。此外,还有哲学和社会伦理的理由。它主要涉及的问题是什么是一个人的优良素质?它是否完全是或主要是由遗传决定的?

(二）个人品质怎样才算优良

人不仅仅是生物学的人,同时也是社会的一员,他不可能不与家庭和社会其他成员相处。而品质从社会伦理学的观点来看,是一个非常复杂的问题。一个人应当具有什么样的品质才算优良呢?如果把身体和精神放在一起讨论,那么常识告诉我们,一个人在某一方面突出,如智商高或会唱歌,并不能说明在其他方面也优秀或超群。因此诺贝尔奖获得者或歌星未必都品质优良;反之,有身体或其他方面缺陷者不乏杰出人物。林肯很可能是马凡氏综合征患者,霍金无法用手书写,两人都是遗传病患者,凡·高、柴可夫斯基都算不得是身心健康的人,而这样的例子还很多。因此,就连最早倡导建立诺贝尔奖金获得者精子库,以实现“人的优化”,本人也是诺贝尔奖获得者的穆勒也意识到:仅仅用高智商作为供精者的标准是远远不够的。于是他又对供精者的品质提出了许多补充要求,包括“友爱、仁慈、慷慨、对生活有审美情趣、充满激情又能自制和性格开朗、道德上的坚贞、思维独立、谦恭乐群、能接受他人公正的批评又能自我批评和改正错误”等。

古今有无符合穆勒要求的,具有上述优秀品质的完人实在难以肯定。历史和现实是,杰出人物,在某一方面优秀的人物确实有,而完人几乎没有。即使是诺贝尔奖获得者,也只能保证他们智力方面没有问题。至于国内“名人精子库”纳入的“名人”,据报载不过是“副高以上职称、硕士以上学位者、明星、高级企业管理人才”等,他们有什么优良的品质,以及什么与之对应的优良基因就更难以说清了。

更有意思的是:为了保证供精者身体和品格上的“优良”不是一时的表面现象,缪勒主张采精后至少要有20年的观察期,才能把他们的精子用于人工授精。这真是“试玉要烧三日满,辨材须待七年期”了。但从生物学的角度看,即便是30岁时获奖,再加上20年,那样的供精者是否又嫌年纪偏高了呢?

在讨论基因与性状,基因与个人品质的关系时,我们不能回避两个具有哲学意义的问题。一是人类的各种性状与品格是否都有一一的对应基因,二是个人品质是怎样形成的。在这里我们暂且将各种躯体的表现称为性状,而把个人精神与心理的表现称为品格。这里有必要区分简单性状和复杂性状。

（三）两类性状

一个个体的各种性状，包括疾病在内，可分为两大类。一类是简单性状，它们是单基因遗传的。另一类是复杂性状，是多基因遗传的。但无论是哪一类，环境因素都起着不同程度的作用。对疾病而言单个基因异常决定的疾病，或单基因疾病，遗传因素非常明显，环境因子的作用相对较小；而在多基因疾病（包括绝大多数肿瘤在内），后天的、环境致病因素的作用就十分突出。因此，在讨论遗传和环境对个体发育、体质状况与疾病发生的影响时，有条件的、相对的遗传决定论是可以接受的，尽管不可忽略环境和后天因素的影响。

但是对于像智力、品格这样非常复杂的非躯体性的表现，如果认为它们也都是基因决定的或先天禀赋的，那就未免太笼统、太简单化了。虽然应当承认个人品格，尤其是智力有一定的躯体的，或生理的，或遗传的基础，但把这一类非常复杂的、精神或心理的表现还原到一个个基因的作用，则是一种机械唯物论的思维方式。它不仅难以做到，也难令人信服。"友爱基因"、"道德基因"、"智力基因"之类不过是简单化了的新闻词语罢了。

基因毕竟是一个遗传学的概念，不应把它滥用来解释复杂的个人品格，更不应滥用来说明社会现象。不排除像个人品格之类的这些复杂现象后面有某些生物学因素在起作用，但后天的，即环境的、家庭的、教育的和社会的因素无疑对个人品格和行为模式的形成起着更大的、决定性作用。迄今对多个基因与环境因素共同决定的躯体性疾病的基因研究尚且没有取得令人信服的结果，更不用说对决定人类品格和高级行为有关的基因了。因此认为缪勒提出的受精者所应具有的各种品质，如认为友爱、仁慈、慷慨等都有对应的基因，而且都能按照简单的孟德尔方式遗传给下一代并发挥作用，至少目前还缺乏令人信服的证据。

五、细胞核的移植和人的克隆

将一个人的二倍体细胞核移入去核的卵细胞，并使之发育成为个体，称为核移植。遗传学上这样的个体在基因型上与亲本供体完全一致，是供体的完全复制品，或克隆。早在20世纪50年代Briggs和King就已完成了蛙的这种无性繁殖，得到了蛙的克隆。作者们当时的目的只是想弄清分化程度不同的胚胎细胞是否具有相同的遗传信息。60年代英国科学家Gurdon等又进一步通过移植已分化的小肠上皮细胞获得了非洲爪蟾的克隆，但哺乳动物的克隆并未实现。因此Rorvik"人的复制"（In His Image，1978）一书在当时只能看作是一本可以乱真，并引起了混乱的科幻读物。这位作者在书中描述了将一位名叫Max的工业家的体细胞移入了一个去核的卵细胞内，然后将早期卵裂球植入一个16岁的孤女的子宫内，从而复制了自身的故事。

其后，中外科学家在利用胚胎细胞核成功地克隆了多种动物，包括鱼、小鼠、牛羊和猴等。

1997年，苏格兰罗斯林研究所的Wilmut用取自6岁成年绵羊乳腺细胞核，通过核移植获得了克隆绵羊多莉。多莉的基因和表型与提供细胞核的绵羊相同。由于这是用已高度分化细胞的核移植达到哺乳动物无性繁殖成功的首例，引起世界舆论的震动是很自然的。因为它表明，从技术上讲，实现人的克隆只是早晚的问题。

克隆技术给生物学研究、动物良种的保存与繁育、生物制药和医学发展产生了巨大的影响。与此同时，克隆技术的应用也引起了争论。有人担心将其用于动植物会导致生物多样性的减弱或消失。但克隆动物的伦理问题并不突出。

然而，当克隆技术用于人类时却可能引起一系列生物学、伦理学和社会问题。迄今公众对这些问题尚未有一致的看法，对接受克隆技术应用的结果还缺乏充分的思想准备。

从生物进化的角度，由低级的无性生殖过渡到有性生殖是一大进步，性和生殖结合也有深刻的进化上的意义。因而一些人认为人的克隆，即人的无性生殖是一种倒退，情爱或性与生殖的分

离是违背自然的。人类为什么要放弃婚姻和有性生殖呢？然而克隆人引起的争议主要还在伦理道德和社会方面。

（一）克隆人的生物学问题

如果一个个体多次提供体细胞核，从而产生大量克隆人，如果不能排除他们之间的婚配的话，那么后果也类似多次采用同一供精者的精细胞作人工授精一样，只是这种近亲婚配可能产生的遗传后果，如隐性遗传病的增加等将更加严重，因为他们的基因组完全一样。

（二）克隆人的家庭伦理学问题

克隆人是体细胞供体的复制品，是无性生殖的产物，那么克隆人与供体是什么关系呢？是供体的又一本体？他的子女？还是兄弟？他（她）与代孕母亲和抚养母亲是什么关系？他与其他用同一供体的细胞核，由同一代孕母亲生育的克隆人之间是什么关系？是否是同父同母的兄弟姐妹？如果代孕母亲不是同一人，他（她）们之间又是什么关系，是否同父异母兄弟？如果一个代育母亲多次代育，而体细胞核的供体不同，这些克隆人之间是什么关系？是同母异父兄弟？如果供体不算作父亲，而算作兄弟，而代孕母亲算作母亲，则情况更加混乱。家庭伦理的混乱可能会影响成员间的相处和克隆儿童幼小心灵的困扰，而相关的法律问题包括亲权，继承权，监护权等迄今都还没有解决。

当然，应当承认，所有这些都是从人们长期习惯了的伦理观念出发来讨论问题的。如果未来克隆人已十分普遍，伦理观念可能改变或逐渐淡化。上述问题也许不会像今天这样令人困惑。

（三）克隆人引起的人权和道德问题

主要的问题是克隆人是否仅是工具和器官组织的来源？克隆人也是人，他理应享有完全的人权，而不应是供研究用或从事危险劳动的工具，更不应是医用器官和组织的来源。因为要是那样的话，我们将回到比奴隶社会更加野蛮的社会。人本体是存在，不是手段和工具，不管他是有性生殖还是无性生殖而来。

（四）治疗性克隆

许多人不赞成克隆人，然而赞同“治疗性克隆”。所谓治疗性克隆（therapeutic cloning），是指将克隆技术用于获得人的早期胚胎，从中分离具有多潜能的干细胞，而后者可分化为皮肤、神经等各种组织，甚至器官，以供治疗和移植使用。因此，尽管许多科学家不赞成“生殖性克隆”（reproductive cloning），即克隆完整的人，但赞成治疗性克隆。如英国科学家已建议政府放宽对人类早期胚胎的克隆研究。

但即便是治疗性克隆，也不能完全回避伦理学问题。尤其是胚胎，因为它是一个完整的人类生命体。在这里我们涉及一些非常困难的问题：人的生命从什么时候开始？从受精？从神经系统发育的某一阶段？从胎儿有感知？其次，在什么阶段或期间他人（包括父母）有权终止其生命，如进行人工流产或作其他处置，以及在什么条件下可以将这种生命体加以利用？对于这些问题，由于宗教、文化和社会背景的差异，人们可能有完全不同的认识。

对治疗性克隆的理解也还不一致。如曾有人认为可以直接克隆人胚胎或无头胎儿，或采用异种克隆办法，将人的供体细胞移植到有关动物的去核卵胞质中，培育出具有人类遗传特性的出生前胎儿，并用于人类临床治疗，但笔者以为，克隆出生前胎儿并用于治疗或研究，尽管采用了动物卵细胞胞质，仍然会引起道德伦理方面的争议。

（五）克隆人问题的立法

近几年来由于不断传出克隆人已获成功的消息，虽然其真实性未能证实，但已引起了国际社会的极大关注，并要求立法加以禁止。然而各国对待克隆人的态度有很大差别，要取得一致和制订禁止克隆人的国际公约还有待时日。从提交给联合国的两项提案可以看出，两项提案均要求禁

止人的生殖性克隆。但一些国家,包括美国在内的提案曾要求禁止所有形式的克隆,而包括比利时、英国、法国、日本、韩国和中国在内的提案则主张区别对待以克隆完整人类个体为目的的生殖性克隆和治疗性克隆,对后者应严加管理,但禁止与否由各国自行决定。双方争论的焦点在于治疗性克隆是否仍有悖伦理,而关键问题仍是人的生命始于何时。美国等国家认为,用于治疗性克隆的胚胎已是新的生命,提取胚胎干细胞后摧毁胚胎与"杀人"无异,而主张准许治疗性克隆的国家则认为,早期胚胎还不是真正意义上的人,用于研究和治疗疾病并不违背伦理,何况其目的是为了治病救人。

幸运的是,近年来干细胞研究的进展为治疗性克隆提供了新的选择。已经发现许多成体组织中也有少量多潜能干细胞,它们可能成为细胞或组织来源,从而避免人类胚胎的利用。如果器官也能这样获得的话,或许可以避免上述一些伦理学问题。

(六)克隆人引起的社会问题

近年来动物的生殖性克隆已取得很大成绩,它也未引发重大的争论。但人的生殖性克隆的需求并未完全消失,忧虑仍然存在。主要忧虑有:①担心性比失调,如果一个社会中多数人偏爱某一性别的话;②担心人类多样性的减弱或消失,而遗传多样性的减弱可能降低人类群体对环境改变的适应性,包括抗病能力等。此外,人们已经长期习惯了多样性的社会。可以肯定,如果足球场上两支球队的22名球员即使全由马拉多纳体细胞发育而来的克隆人组成,那样的比赛也会十分乏味;③人们担心克隆技术会激活纳粹主义者或妄自尊大的优生思潮,并用于不正当的目的。例如,认为某一民族、家族或个人特别优秀而将其是大量克隆。又如,克隆身体素质特别优良、勇猛而又绝对服从的人作为战士,而政治家担心这样的队伍如果落入独裁者手中将是民主制度的终结。再如,克隆大量球星并在少年足球学校加以训练,待长成后出售给足球俱乐部等。当然,在现阶段,要做到这一点还有困难,因为许多野心家或商人希望赋予克隆人的品质绝大多数都是多基因决定的,而诸如勇猛、服从或球艺高超之类的品质,更远非仅由基因决定的。特别应强调的是:克隆人由于后天成长环境和教育的不同,他只是供体遗传上的复制品,但其他方面不会彼此完全相同。

尽管人们有上述各种忧虑,人的克隆仍然有多方面的诱惑力。归纳起来,它来自如下几个方面:①对永生不灭的追求,因为可能有人把克隆人视为自己的又一本体或生命的绵延继续;②对重新生活一遍的向往,包括在另一时代或另一环境下重新生活的向往,尽管克隆人与供体的"同一性"(identity)可能不同;③希望重新获得丧失的子女或亲属;④希望重新获得被崇拜的政治或宗教领袖。至于为了不正当的商业、政治和军事目的而克隆人,前面已经提到,不再赘述。

从有关人的克隆的讨论可以看出,遗传工程和细胞工程技术的发展既可以为人类造福,也可能引起各种严重的问题。正确的态度是:①继续发展干细胞技术和治疗性克隆技术,并促进其在已达成共识,有利于人类的大多数领域内应用;②加强医学伦理学,特别是遗传伦理学和相关社会问题的研究,尽快在一些方面达成共识,制定相应的伦理道德准则和法规,并关注未来可能出现的新问题,从而促进遗传新技术的发展与应用,同时以避免其滥用。

Summary

Genetic ethics is the knowledge of the moral and ethical regulations which should be observed during genetic service in medicine.As a kind of medical practice genetic service should follow all general principles of the medical ethics.At present the genetic service mainly consists of genetic counseling,genetic testing,gene therapy and genetic service in assisted reproduction including artificial insemination, *in vitro* fertilization,embryo or nucleus transfer etc.

During the genetic counseling the most important is to understand the state of mind or mentality of the counselee, and to follow the principles of respect for his privacy, autonomy, informed consensus, and to avoid any directive suggestions. The same principles plus nonmaleficence, beneficence, should be observed in genetic testing or diagnosis. But, whether to do the test or not in case of late-onset adult incurable inheritable disease, is sometimes a difficult choice because people do not know which is better for the patients. The same situation is whether to test the children who even cannot fulfill their right of informed consensus and, if the disease is incurable, the testing will be of no benefit to them and may even make them suffer from discrimination.

The safety and effectiveness of gene therapy for most diseases and cancers should be carefully evaluated, particularly when germ cell gene therapy is applied. Besides, the indications and requirements for gene therapy should be observed. It is still debatable whether we should "improve" our genetic quality for any purpose other than disease treatment.

More ethical issues arise from the use of assisted reproduction techniques (ART) such as artificial insemination, in vitro fertilization and embryo or somatic cell nuclear transplantation. Because it provides the opportunity to adopt germ cells, embryos and somatic cell nucleus of other people and also the possibility of selection of genetic materials, thus arouses many family ethical and social issues. Therefore, it is important to emphasize the study on genetic ethics and related issues while we develop and apply any new techniques in order to reach a consensus on its medical application and to avoid the possible abuse or misuse of novel genetic technology and service.

（张思仲）

参考文献

1. 邱仁宗. 生命伦理学. 上海：上海人民出版社，1987
2. Wertz DC, Fletcher JC. Reilly P. Ethnics and medical genetics in the United States. A national survey. Am J. Med. Genet 1988, 29: 815-827
3. 奥柏瑞·米兰斯基. 别让遗传决定你的命运. 台北：商周出版社，2000
4. 范晔. 后汉书. 北京：中华书局，2007
5. W. 富尔曼，F. 沃格尔. 遗传咨询. 上海：上海翻译出版公司，1983
6. 倪慧芳，刘次全，邱仁宗. 21 世纪生命伦理学难题. 北京：高等教育出版社，2000
7. Harper P. 分子遗传学的社会和伦理问题——西方的问题及其与中国的关系（张思仲译）. 中华医学遗传学杂志，1993, 10(4): 193-197
8. Tjio JH, Levan A. The metaphase chromosome number in man. Hereditas, 1956, 42: 1
9. Caspersson T. The 24 fluorescence patterns of human metaphase dhromosomes-distinguishing characters and variability. Hereditas, 1971, 67: 89
10. Harper PS, O' Bren T, Murray JM, et al. The use of linked DNA polymorphisms for genotype prediction in families with Duchenne muscular dystrophy. J. Med. Genet, 1983, 20: 252
11. Deech R. Guide to donor insemination and IVF clinics. Hum Reprod. , 1996, 11: 1863-1370
12. Rolvik D. In His Image: The Cloning of a Man. Philadelphia: J. B. Lippincott Co, 1978
13. Graver K L, and Graver B. The human genome project and eugenics concerns. Am J. Hum Genet, 1994, 54: 148-158
14. Report of a WHO Meeting on Ethical Issues in Medical Genetics: Proposed International Guidelines on Ethical Issues in

Medical Genetics and Genetic Services. Geneva, 15-16 December 1997. World Health Organization, 1998

15. Mahowald MB, Verp MS, Anderson RR. Genetic counseling: clinical and ethnical challenges. Annu. Rev. Genet., 1998, 32: 547-559
16. HUGO Ethics Committee: 1999, Statement on Cloning, Brisbane, March 1999
17. Veach PM, LeRoy BS, Bartels DM. Facilitating the Genetic Counseling Process · Chapter11 Behaving Ethically. Springer-Verlag, 2003, 222-241

汉英索引

A

B

C

D

E

F

G

H

I

J

K

L

M

N

P

Q

R

S

T

W

Z

英汉索引

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y